

Una Versión Reducida de la Escala de Actitudes hacia la Estadística

A Short Version of the Scale to Measure Attitudes Towards Statistics

Ana María Ruiz-Ruano García¹ y Jorge López Puga²

¹ Departamento de Psicología Evolutiva y de la Educación, Facultad de Ciencias de la Educación,
Universidad de Granada

² Departamento de Personalidad, Evaluación y Tratamiento Psicológico, Facultad de Psicología,
Universidad de Granada

La estadística es una materia transversal a un amplio espectro de áreas de conocimiento. La relevancia de esta materia en los contenidos educativos contrasta con la afinidad que el estudiantado tiene hacia ella. Una actitud negativa hacia la estadística puede explicar el bajo rendimiento que el alumnado universitario ha mostrado históricamente en esta asignatura. Una muestra no probabilística obtenida por medio de muestreo de conveniencia de 384 estudiantes de una universidad privada española (Región de Murcia) tomó parte en este estudio y respondió en dos ocasiones, antes del inicio del curso y al final de este, a la escala de Actitudes hacia la Estadística de Auzmendi (1992). Se analizaron dos modelos de la escala utilizando análisis paralelo y análisis factoriales confirmatorios, uno incluyendo todos los ítems y otro en el que el número de ítems se redujo a 15. Se observó que la versión reducida de la escala mostró buenas propiedades psicométricas en términos de fiabilidad, en sus aspectos de consistencia interna y temporal, y a nivel de la estructura factorial. Las puntuaciones de la versión reducida de la escala correlacionan positiva y significativamente con la nota en un examen parcial y en la nota final de la asignatura de análisis de datos. Esta versión reducida puede considerarse más parsimoniosa que la original con utilidad educativa aplicada y para la investigación sobre actitudes hacia la estadística.

Palabras clave: actitud, estadística, medida, fiabilidad, validez

Statistics is a transversal matter in a wide range of knowledge areas. Although statistics is critically important for college students, most of them struggle when studying. Negative attitudes towards statistics can explain the weak academic performance that has been historically observed in college students. A non-probabilistic sample of 384 undergraduate students from a Spanish private university (Región de Murcia) recruited by a convenience sampling answered the Attitudes towards Statistics (Auzmendi, 1992) scale twice (at the beginning and at the end of the course). Two models were analyzed by using parallel analysis and confirmatory factorial analyses, the original version and a short version including only 15 items. The short version scores correlated positively and significantly with students grades in the mid-term and at the end of the course on data analysis. The short version showed acceptable reliability (in terms of internal consistency and temporal stability) and factorial structure. As a result, the short version is considered more parsimonious than the original version and optimal to use in educative context and for research purposes.

Keywords: attitude, statistics, measurement, reliability, validity

Los conocimientos en el área de la estadística pueden ser considerados transversales o básicos para muchos ámbitos de estudio. De hecho, ha llegado a proponerse que la "alfabetización estadística" es la clave para un futuro próspero de la humanidad (Bond, 2009). La relevancia de la estadística es clara en el ámbito

Ana María Ruiz-Ruano García  <https://orcid.org/0000-0002-7260-0588>

Jorge López Puga  <https://orcid.org/0000-0003-0693-0092>

Para dar transparencia a este trabajo, algunos resultados no incluidos en el presente artículo, como el código fuente para R, utilizado para analizar los datos, y los datos brutos de la investigación pueden encontrarse en el siguiente proyecto de la Open Science Framework: <https://osf.io/6xrgk/> Una breve contextualización de la información contenida en el proyecto está disponible en <https://osf.io/6xrgk/wiki/home/>

La correspondencia relativa a este artículo debe ser dirigida a Jorge López Puga, Departamento de Personalidad, Evaluación y Tratamiento Psicológico, Universidad de Granada, Facultad de Psicología, Campus de la Cartuja s/n, 18071, Granada, España. Email: jlpuja@ugr.es

de las ciencias sociales, contexto en el que surgió esta disciplina de estudio, pero también es esencial en estos días para las ciencias de la salud, las ciencias naturales e, incluso, para la física, la química o las telecomunicaciones (Rodríguez Parrondo, 2019). Por ello, la mayor parte de las titulaciones universitarias incluyen en sus programas educativos alguna asignatura sobre estadística o análisis de datos. Sin embargo, algunos estudiantes experimentan problemas cuando se enfrentan al estudio de la estadística. Por algún motivo, el trasfondo matemático que implica esta disciplina, así como la lógica que en ella subyace, genera cierta aversión en los estudiantes. Algunos autores sugieren que la estadística genera cierto tipo de ansiedad que es particularmente específica y que puede ser considerada como una de las más perniciosas en el contexto educativo (Chew et al., 2017). La ansiedad que los estudiantes experimentan hacia la estadística parece estar derivada del componente matemático que hay en la estadística (Estrada et al., 2013). Dado que la ansiedad a las matemáticas ha sido relacionada con bajos niveles de atención y memoria de trabajo (Hartwright et al., 2018), minimizar esta forma de ansiedad en el contexto educativo es crucial para optimizar el proceso de enseñanza-aprendizaje.

Los problemas con la estadística no solo afectan a los estudiantes universitarios o de secundaria, sino que la comunidad científica también experimenta ciertas complicaciones cuando trata de usar e interpretar conceptos estadísticos. El uso y la interpretación inapropiada de los estadísticos inferenciales comúnmente producidos en los contrastes de hipótesis es un tema que viene siendo discutido desde los años 60 del pasado siglo (véase, por ejemplo, Bakan, 1966). Lejos de haberse superado a lo largo de los últimos años, parece, como señalan Pascual Llobell et al. (2000), que se trata de un tipo de problema recurrente que se vuelve a producir cíclicamente. Por ejemplo, en una investigación realizada por Haller y Krauss (2002) se observó que tanto los alumnos universitarios como sus profesores tenían concepciones erróneas del contraste estadístico de hipótesis y del valor p de un test estadístico. Por ello, se han propuesto numerosas guías y/o directrices destinadas a promover un uso apropiado de este tipo de estadísticos (véase, por ejemplo, Aczel et al., 2020; Altman & Krzywinski, 2017; Gao, 2020; Kruschke, 2021; Nuzzo, 2014; Orlitzky, 2012; Ruiz-Ruano García & Puga, 2020; Wasserstein & Lazar, 2016). En cualquiera de los casos, todavía no existe consenso en determinar si este tipo de problemas se debe a deficiencias comunicativas por parte de la comunidad científica, si es un problema de comprensión o si podría explicarse por malas prácticas de investigación (Baker, 2016).

Algunos profesionales educativos sostienen que el problema tiene su base en el contexto formativo y que las soluciones han de provenir de ese ámbito (Bishop, 2020). De manera general, los alumnos universitarios no suelen recibir con entusiasmo los contenidos relacionados con la estadística. Muchos alumnos tienden a acoger los contenidos de estadística con recelo y tienden a considerarlos poco útiles para el desarrollo de su carrera académica (Acón Araya & Salazar Ramos, 2020; Ruiz de Miguel, 2015). Todo ello tendría como consecuencia un bajo rendimiento académico en las asignaturas universitarias sobre estadística (Bologna & Vaiman, 2013). Este fenómeno es especialmente llamativo en los primeros cursos universitarios en los que se ha observado hasta un 80% de fracaso académico en las asignaturas relacionadas con el análisis de datos estadístico (Ponsoda, 1990).

Ante esta tesitura, numerosos docentes han emprendido acciones encaminadas a diseñar entornos de enseñanza-aprendizaje que optimicen los procesos de asimilación de competencias y contenidos estadísticos (véase, por ejemplo, Bates, 1991; Gorvine & Smith, 2015; Hartnett, 2013; Karpik, 2011; León, 1996; López Puga, 2014, Ruiz-Ruano García, Palazón Pérez de los Cobos et al., 2018). No obstante, no resulta fácil para el profesorado —incluso algunas veces, desalentador—, generar métodos de enseñanza amigables, dada la heterogeneidad de sensibilidades que generan los contenidos estadísticos en los alumnos (Vilà Baños & Rubio Hurtado, 2016). Las iniciativas que parecen haber mostrado más resultados positivos son aquellas destinadas a minimizar la sensación de amenaza que la estadística genera en los alumnos, así como la presentación lenta y progresiva de sus contenidos (Bologna & Vaiman, 2013). Gorvine y Smith (2015) pusieron en marcha un programa destinado a reducir la ansiedad hacia la estadística basado en el trabajo colaborativo. Sus resultados mostraron una reducción en los niveles de ansiedad y un aumento en el aprendizaje. Desde el punto de vista curricular, Barron y Apple (2014) observaron que es más productivo presentar cursos de estadística integrados con componentes de diseños de investigación o metodología científica. Por su parte, Ruiz-Ruano, Palazón et al. (2018) y Tintle et al. (2015) proponen potenciar el uso de tareas de simulación para mejorar el pensamiento estadístico en estudiantes a nivel universitario.

Dado que la actitud es considerada como uno de los factores que condicionan el comportamiento (Ajzen & Fishbein, 2005; Allport, 1935; Fishbein & Ajzen, 1975; Morales et al., 1994), conocer las actitudes hacia la estadística puede servir para predecir el comportamiento y, potencialmente, orientarlo. Por tanto, conocer las actitudes que el estudiantado tiene hacia la estadística en un contexto determinado puede ser de utilidad

para desplegar las acciones pertinentes destinadas a optimizar el proceso de enseñanza-aprendizaje. Estudios previos han mostrado que la actitud del profesorado hacia la estadística también condiciona la actitud del alumnado hacia esta disciplina (Estrada et al., 2004, 2013). Este fenómeno es crítico, por ejemplo, cuando se aborda la probabilidad, ya que es uno de los tópicos más importantes de esta disciplina académica (Estrada & Batanero, 2015).

Existen numerosas escalas que han sido diseñadas para medir las actitudes hacia la estadística. En la revisión sistemática de Nolan et al. (2012) se hace un repaso histórico y conceptual reciente sobre estas escalas, destacando las características de cada una de ellas. Una de estas escalas fue diseñada por Auzmendi (1992).

El objetivo de este trabajo fue valorar la idoneidad psicométrica de una versión recortada de esta escala. Esta iniciativa está justificada por varios motivos. En primer lugar, porque en estudios previos se ha observado que los ítems de la escala primigenia podrían agruparse de una forma diferente a como se planteó en el estudio original (véase, por ejemplo, Darias Morales, 2000). Por otro lado, en un estudio reciente llevado a cabo por Ruiz-Ruano García, Puga et al. (2018) se observó que un conjunto de expertos mundiales en estadística y metodología de las ciencias del comportamiento agrupó los ítems de la escala de tal modo que algunos de ellos fueron considerados redundantes. Por último, si la escala reducida que se iba a testar mostrase mejores propiedades psicométricas, en comparación con la versión original, podría disponerse de una herramienta más parsimoniosa para evaluar las actitudes hacia la estadística, tanto en contextos educativos como de investigación. Así, en contextos educativos, esta herramienta podría ser de utilidad para llevar a cabo análisis diagnósticos o de cribado orientados a optimizar los procesos de enseñanza-aprendizaje, mientras que en los contextos de investigación podría ser de utilidad para identificar los factores que afectan a los diferentes componentes de la actitud hacia la estadística.

Método

Participantes

La muestra fue seleccionada usando un muestreo por conveniencia y estuvo compuesta por 384 estudiantes del Grado en Psicología de la asignatura Análisis de Datos, Universidad Católica de Murcia, España. El 22,9% eran hombres ($n = 88$) y el 77,1%, mujeres ($n = 296$) con una edad comprendida entre los 18 y los 52 años ($M = 20,76$, $DT = 3,63$). Con relación a los estudios de ingreso a la carrera, 41,4% había cursado bachillerato de Ciencias Sociales, el 38%, bachillerato de Ciencias de la Salud, el 12%, bachillerato de Humanidades, el 4,4%, bachillerato Tecnológico, el 0,5%, bachillerato Artístico y el 3,6%, otros estudios. En respuesta a la pregunta de si habían estudiado previamente estadística, el 42,7% contestó que no. La submuestra destinada a estimar la estabilidad temporal de la escala estuvo compuesta por 170 estudiantes del Grado en Psicología de la asignatura Análisis de Datos, ya que no todos los estudiantes iniciales participaron en la segunda fase del estudio. El 17,5% fueron hombres ($n = 30$) y el 82,4%, mujeres ($n = 141$), con una media de edad de 20,99 años (rango = 18-52, $DT = 4,4$). Con relación a los estudios de ingreso a la carrera, 49,7% había cursado bachillerato de Ciencias Sociales, el 36,8%, bachillerato de Ciencias de la Salud, el 7%, bachillerato de Humanidades, el 3,5%, bachillerato Tecnológico y el 2,9%, otros estudios. En respuesta a la pregunta de si habían estudiado previamente estadística, el 39,2% contestó que no.

Instrumento

La escala de actitudes hacia la estadística desarrollada por Auzmendi (1992) fue utilizada en este estudio. Esta prueba consta de 25 ítems que se agrupan en cinco componentes o dimensiones de la actitud: Valor-Utilidad percibida de la estadística, Agrado hacia la estadística, Ansiedad suscitada por la estadística o su estudio, Motivación hacia el estudio o el uso de la estadística y Seguridad-Confianza frente al uso o estudio de la estadística. Cada dimensión está integrada por cinco ítems que se puntúan en una escala tipo Likert de cinco alternativas en las que la persona tiene que indicar su grado de acuerdo con el ítem (el valor 1 indica *total desacuerdo* con lo expresado por el ítem, mientras que el valor 5 indica *total acuerdo*). Diez de los ítems de la escala (los que se incluyen en las dimensiones de Ansiedad y Motivación) son considerados inversos y han de ser recodificados antes de calcular la puntuación total de actitud hacia la estadística. Una vez recodificados, los ítems se agregan por suma simple y se obtiene una puntuación total de actitud hacia la estadística que oscila entre 25 y 125. A mayor puntuación en la escala se considera que existe una mayor

actitud positiva hacia la estadística en la persona evaluada. Las puntuaciones para cada una de las dimensiones de la escala también pueden ser calculadas.

La estructura factorial inicial de la escala explicaba el 60,7% de la varianza (Auzmendi, 1992), aunque estudios posteriores observaron una agrupación diferencial de los ítems (Darias Morales, 2000). La consistencia interna de la escala, evaluada con el coeficiente alfa de Cronbach, fue de 0,93 en el estudio original, mientras que en el estudio de Darias Morales (2000) fue de 0,90. Las consistencias internas de las dimensiones de la escala han oscilado entre 0,49 y 0,91. En un estudio reciente en el que se utilizó el análisis de redes (Ruiz-Ruano García et al., 2020) se observó que, en concordancia con el estudio original de Auzmendi (1992), los ítems de Ansiedad se organizaron en una red más cohesionada y distante del resto de ítems, mientras que los ítems de Utilidad y Agrado mostraban redes más fusionadas.

La versión reducida de la escala estuvo compuesta por 15 ítems (ver Tabla 1). Para seleccionar los ítems que formaron parte de esta versión reducida se tuvo en cuenta el análisis y la propuesta realizada por Ruiz-Ruano García, Puga et al. (2018). En concreto, dos ítems fueron eliminados de cada dimensión, siguiendo las sugerencias hechas por una muestra de expertos en análisis estadístico de datos y metodología de las ciencias del comportamiento. Los expertos (14) formaban parte de un panel de 148 profesores e investigadores que habían publicado artículos recientes sobre la temática en revistas indexadas en la *Journal Citations Reports*. Los expertos indicaron, según su criterio, a qué dimensión pertenecía cada uno de los ítems de la escala. Aquellos ítems en los que los expertos mostraron un menor grado de acuerdo fueron eliminados de la versión reducida de la escala. Así, los ítems 2, 3, 5, 6, 10, 12, 13, 19, 20 y 24 de la escala original fueron eliminados en la versión reducida. De esta manera, cada dimensión estaría representada por tres ítems y seis de los que integran el test serían inversos.

Tabla 1

Versión Reducida de la Escala de Actitudes hacia la Estadística de Auzmendi (1992)

PIR	POI	Ítem	DIM
1	1	Considero la estadística como una materia muy necesaria en la carrera	VAL-UTI
2	4	El utilizar la estadística es una diversión para mí	AGR
3	7	* La estadística es una de las asignaturas que más temo	ANS
4	8	Tengo confianza en mí mismo/a cuando me enfrento a un problema de estadística	SEG-CON
5	9	Me divierte el hablar con otros de estadística	AGR
6	11	Saber utilizar la estadística incrementaría mis posibilidades de trabajo	VAL-UTI
7	14	La estadística es agradable y estimulante para mí	AGR
8	15	* Espero tener que utilizar poco la estadística en mi vida profesional	MOT
9	16	* Para el desarrollo profesional de nuestra carrera considero que existen otras asignaturas más importantes que la estadística	MOT
10	17	* Trabajar con la estadística hace que me sienta muy nervioso/a	ANS
11	18	No me altero cuando tengo que trabajar en problemas de estadística	SEG-CON
12	21	Para el desarrollo profesional de mi carrera una de las asignaturas más importantes que ha de estudiarse es la estadística	VAL-UTI
13	22	* La estadística hace que me sienta incómodo/a y nervioso/a	ANS
14	23	Si me lo propusiera, creo que llegaría a dominar bien la estadística	SEG-CON
15	25	* La materia que se imparte en las clases de estadística es muy poco interesante	MOT

Nota. PIR: posición del ítem en la escala reducida, POI: posición del ítem en la escala original, DIM: dimensión, VAL-UTI: valor utilidad, AGR: agrado, ANS: ansiedad, SEG-CON: seguridad confianza, MOT: motivación. Los ítems precedidos por asterisco son inversos.

Procedimiento

La escala original fue administrada en formato electrónico la primera y la última semana de clase en la asignatura de Análisis de Datos perteneciente al Grado en Psicología. La recogida de datos se llevó a cabo a lo largo de cuatro cursos consecutivos (2016, 2017, 2018 y 2019) en los meses de octubre (primera medida) y diciembre (segunda medida) de cada curso académico. Se tomaron dos medidas de las actitudes hacia la estadística con el objetivo de estimar la estabilidad temporal de la versión reducida de la escala. La cumplimentación de la escala fue voluntaria, aunque se instaba al alumnado a contestarla por considerarse una herramienta de ayuda para monitorizar el proceso de enseñanza-aprendizaje. Las respuestas a los ítems fueron totalmente anónimas y, para garantizar dicho anonimato, cada uno de los participantes introdujo un identificador autogenerado que solo él/ella conocía y que era necesario para poder vincular sus datos de la primera administración del formulario y la segunda de retest. También se pidió a los alumnos que indicasen la nota que habían obtenido en un examen parcial que se realizó cuando se había tratado el 50% de los contenidos del curso. Además, cada alumno proporcionó su nota final en la asignatura tras la finalización de esta. No se proporcionó ningún tipo de compensación por la participación, excepto el agradecimiento al final de la cumplimentación de cada una de las escalas.

El estudio fue desarrollado tras recibir la aprobación del Comité de Ética de la Universidad Católica de Murcia, al amparo del proyecto de innovación docente titulado “Mejora de las actitudes hacia la estadística en alumnos de psicología”, con código de registro 2018/0399.

Análisis de Datos

La estimación de la fiabilidad, en su aspecto de consistencia interna, se realizó utilizando tanto el coeficiente alfa de Cronbach (α) como el coeficiente omega (ω), dado que este segundo asume ciertas restricciones sobre los datos muestrales que son más plausibles en la práctica (Dunn et al., 2014; Kelley & Cheng, 2012). Para comparar la bondad de ajuste de las versiones original y reducida de la escala se llevaron a cabo análisis factoriales confirmatorios, usando el método de estimación de máxima verosimilitud. Para valorar la bondad de ajuste diferencial de ambas versiones de la escala se tomaron como referencia las directrices y puntos de corte ofrecidos por Ruiz et al. (2010). Se consideró que los índices de bondad de ajuste (índice comparativo de Bentler-Bonett, índice de Tucker-Lewis, índice no normalizado de Bentler-Bonett, índice normalizado de Bentler-Bonett, índice normalizado de parsimonia, índice relativo de Bollen, índice incremental de Bollen e índice de no centralidad relativa) ponían de manifiesto mejor ajuste cuando estuvieron más próximos a uno. Por su parte, se consideró que los criterios de información o indicadores de parsimonia (logaritmo de la verosimilitud, criterio de información Akaike y criterio de información Bayesiano) así como las estimaciones del error (error cuadrático medio de aproximación, intervalo de confianza al 90% para el error cuadrático medio de aproximación y error cuadrático medio residual estandarizado) denotaban mejor ajuste cuanto más próximos estuvieron a cero. Se realizaron análisis paralelos para la versión original y reducida de la escala con base en las directrices sugeridas por Falissard (2011). Para ello se extrajeron 20 muestras aleatorias a partir de los datos y se superpusieron estas simulaciones a los datos observados en los correspondientes gráficos de sedimentación. Las variables relativas a la nota obtenida a mitad de curso y al final de este fueron utilizadas para valorar el grado en que las puntuaciones de la escala covarían con el rendimiento de los alumnos. Las notas de la asignatura, intermedias y finales, se asignaron teniendo en cuenta los criterios del sistema de calificaciones en las titulaciones universitarias de carácter oficial de España (Boletín Oficial del Estado número 224, Real Decreto 1125/2003 de 5 de septiembre). Por tanto, se valoraron las notas con un valor comprendido entre cero y diez más una cifra decimal. Además, en consonancia con ese sistema nacional de notas, la nota final de la asignatura se categorizó usando las siguientes categorías: una puntuación comprendida entre 0 y 4,9 se consideró como un “suspenso”, una nota comprendida entre 5,0 y 6,9 se consideró como un “aprobado”, una nota comprendida entre 7,0 y 8,9 se consideró como un “notable” y una puntuación comprendida entre 9,0 y 10 se consideró como un “sobresaliente”. La mención de “matrícula de honor” se asignó a las personas que habían obtenido una nota igual o superior a 9,0 cuando el rendimiento de la persona destacó por su excelencia según el criterio personal del profesorado. Por normativa, la cantidad de matrículas de honor nunca superó el 5% del número de personas matriculadas en la asignatura.

Además de la estimación clásica de los intervalos de confianza y del valor p clásico para los contrastes de hipótesis relativos a las correlaciones estimadas, se calcularon intervalos de credibilidad bayesianos y los correspondientes factores de Bayes favorables a las hipótesis alternativas (FB_{10}) o al modelo estadístico testado (FB_M). Los intervalos de confianza para la correlación se estimaron considerando una distribución previa no informativa (Aczel et al., 2020), siguiendo la recomendación de Rouder et al. (2009), con un factor de escalado $r = 1$. Se optó por este factor de escalado, porque los estudios recientes de simulación han mostrado que el comportamiento de este tipo de distribución previa tiene un funcionamiento aceptable en un amplio abanico de situaciones (Jeon & De Boeck, 2017). Además, como sugiere Kruschke (2021), se llevaron a cabo análisis de sensibilidad destinados a valorar el grado en que los factores de Bayes fluctuaban en función de la elección de una u otra función de escalado previa. En ningún caso se observó que la modificación del escalado previo modificase de manera relevante los factores de Bayes estimados. Para la estimación de los modelos de análisis de varianza unifactorial se utilizó una distribución previa equiprobable para el modelo de efectos fijos y para el modelo nulo con probabilidad de 0,5 para ambos modelos. Los factores de Bayes se interpretaron atendiendo a las categorías indicadas por Jeffreys (1948) y asumiendo la naturaleza continua de este estadístico, así como considerando todas las cautelas relativas al uso de este tipo de estadísticos inferenciales (Ruiz-Ruano García & Puga, 2020). Por tanto, los factores de Bayes comprendidos entre 1 y 3 fueron considerados como una evidencia de la hipótesis alternativa que únicamente merece una simple mención (Kass & Raftery, 1995), cuando estuvieron comprendidos entre 3 y 10, como evidencias substanciales, cuando estuvieron comprendidos entre 10 y 30, como evidencias fuertes de la hipótesis alternativa, cuando estuvieron comprendidos entre 30 y 100, como evidencias muy fuertes y cuando los factores de Bayes superaron el valor 100, como evidencias decisivas en favor de la hipótesis alternativa.

Para analizar los datos se utilizó la versión 0.13.1 de JASP (JASP Team, 2020). Para estimar el modelo bifactorial se utilizó la versión 0.6-9 del paquete *lavaan* (Rosseel, 2012) para el software de análisis estadístico R (versión 4.1.1; R Core Team, 2021).

Resultados

En la Tabla 2 aparecen los estadísticos de bondad de ajuste para las versiones original y reducida de la escala. Como se puede apreciar, la bondad de ajuste de la versión reducida obtiene mejores resultados que la versión original. Tanto los índices de ajuste normalizados como los no normalizados son mayores para la escala reducida y sus valores reflejan un buen ajuste. Además, el menor valor del PNFI para la versión reducida sugiere que este formato del test es más parsimonioso que el de la versión original. En cuanto a los estadísticos referidos a criterios de información, también se puede observar que la versión reducida obtiene índices sistemáticamente menores, en comparación con la versión original. Por último, los estadísticos referidos a la estimación del error también muestran que la versión reducida está consistentemente asociada a menores tasas de error.

Tabla 2

Índices de Bondad de Ajuste de la Versión Original (Auzmendi, 1992) y Reducida de la Escala de Actitudes hacia la Estadística

Estadístico	VO	VR	VR _{bifac}
<i>Índices de ajuste</i>			
χ^2 del modelo base	4793,85	2516,90	2516,90
Grados de libertad del modelo base	300	105	105
χ^2 del modelo	672,98	184,38	330,15
Grados de libertad del modelo	265	80	90
χ^2/df	2,54	2,30	3,67
Índice comparativo de Bentler-Bonett (CFI)	0,91	0,96	0,90
Índice de Tucker-Lewis (TLI)	0,90	0,94	0,88
Índice no normalizado de Bentler-Bonett (NNFI)	0,90	0,94	0,84
Índice normalizado de Bentler-Bonett (NFI)	0,86	0,93	0,90
Índice normalizado de parsimonia (PNFI)	0,76	0,71	0,75
Índice relativo de Bollen (RFI)	0,84	0,90	0,85
Índice incremental de Bollen (IFI)	0,91	0,96	0,90
Índice de no centralidad relativa (RNI)	0,91	0,96	0,90

(continúa)

Tabla 2 (Conclusión)

Índices de Bondad de Ajuste de la Versión Original (Auzmendi, 1992) y Reducida de la Escala de Actitudes hacia la Estadística

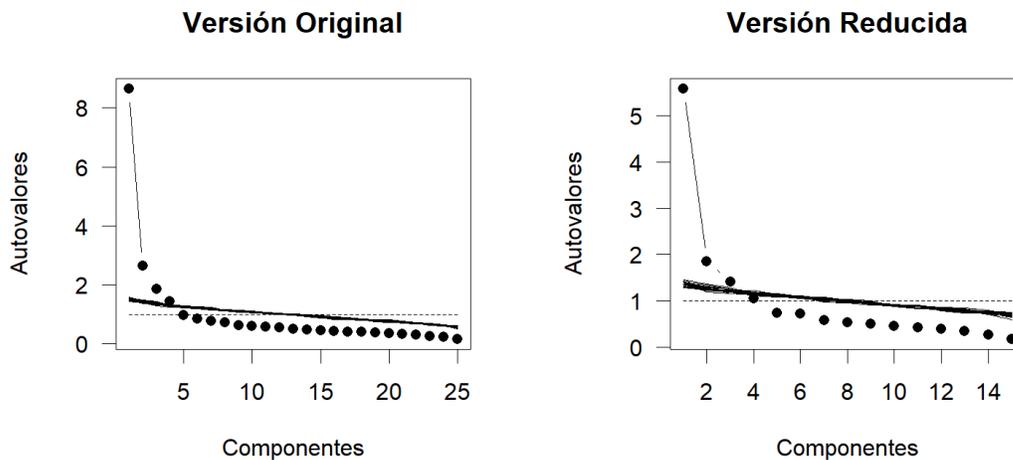
Estadístico	VO	VR	VR _{bifac}
<i>Criterios de información</i>			
Logaritmo de la verosimilitud	-12285,53	-7430,32	-7503,20
Criterio de información Akaike (AIC)	24691,07	14940,63	15066,40
Criterio de información Bayesiano (BIC)	24928,11	15098,66	15184,92
<i>Estimación del error</i>			
Error cuadrático medio de aproximación (RMSEA)	0,063	0,058	0,083
Intervalo de confianza al 90% para RMSEA	[0,057, 0,069]	[0,047, 0,069]	[0,074, 0,093]
Error cuadrático medio residual estandarizado (SRMR)	0,059	0,051	0,094

Nota. VO: versión original de la escala, VR: versión reducida de la escala, VR_{bifac}: modelo bifactorial para la versión reducida.

Sin embargo, los análisis paralelos realizados sugieren que las estructuras factoriales de la versión original y su análoga reducida podrían ser diferentes a lo que se propone teóricamente. Como se puede apreciar en la Figura 1, el análisis paralelo sugiere que una estructura de cuatro factores podría ser plausible, aunque, aplicando el criterio de Keyser, sí sería verosímil una estructura factorial de cinco factores. En lo que respecta a la versión reducida, el análisis paralelo sugiere que una solución con tres factores podría ser aceptable, mientras que el criterio de Keyser apunta a que sería suficiente con una solución de cuatro factores. Por su parte, el análisis bifactorial no implica una mejora sustancial en los índices de bondad de ajuste. Además, como se observa en la Tabla 1, las tasas de error de este modelo son mayores, comparadas con las del modelo de cinco factores.

Figura 1

Gráficos de Sedimentación para los Análisis Paralelos de la Versión Original y Reducida de la Escala



Nota. La línea horizontal discontinua muestra el punto de corte correspondiente a la regla de Keyser para valorar la dimensionalidad.

Como se muestra en la Tabla 3, los índices de consistencia interna para la versión reducida de la escala son muy similares a los índices observados para la versión original. La reducción en la consistencia interna que se nota para la puntuación total de escala en su versión reducida, así como la que se advierte para cada una de las dimensiones, puede explicarse por la disminución en el número de ítems. También se aprecia que,

tanto para la versión original como para la versión reducida, la dimensión con mayor consistencia interna es Ansiedad, mientras que la que muestra una consistencia interna menor es Motivación.

Tabla 3
Fiabilidad de la Escala Original y Reducida y de sus Dimensiones

Escala y dimensión	Original		Reducida	
	ω	A	ω	α
Total	0,92	0,91 [0,90, 0,93]	0,88	0,87 [0,86, 0,89]
Ansiedad	0,89	0,89 (0,87, 0,91)	0,89	0,89 (0,86, 0,90)
Motivación	0,67	0,68 (0,63, 0,83)	0,61	0,61 (0,53, 0,67)
Agrado	0,86	0,86 (0,83, 0,88)	0,83	0,82 (0,79, 0,85)
Seguridad	0,85	0,85 (0,82, 0,87)	0,74	0,73 (0,68, 0,77)
Utilidad	0,74	0,73 (0,68, 0,77)	0,70	0,69 (0,63, 0,74)

Nota. Entre corchetes aparecen los intervalos de confianza al 95%.

En lo que respecta a la fiabilidad de la escala en su aspecto de estabilidad temporal (Tabla 4), se puede ver que, aunque el coeficiente de fiabilidad test-retest es dos centésimas menor en la escala reducida, la estabilidad temporal está más equilibrada en las subescalas de esta versión. Concretamente, la estabilidad temporal es mayor en las dimensiones Ansiedad y Motivación, aunque es menor en las otras tres restantes. Es destacable que, atendiendo al factor de Bayes, la evidencia muestral en favor de la estabilidad temporal para la dimensión Motivación puede ser considerada únicamente fuerte para la versión original mientras que puede ser considerada decisiva para la versión reducida. Es decir, mientras que el Factor de Bayes es ligeramente mayor que 20 cuando se estima la correlación entre la primera y segunda administración de la escala original, el Factor de Bayes obtenido es considerablemente mayor que 100 para la correlación de las puntuaciones inicial y final obtenidas con la versión reducida de la escala. Esto sugiere que los datos muestrales son más congruentes con la hipótesis de estabilidad temporal en la dimensión de Motivación cuando se usa la versión reducida de la escala.

Tabla 4
Correlación entre las dos Administraciones de los Dos Formatos de la Escala en cada Una de las Fases de Recolección de Información

Escala y dimensión	R	95% IC	FB_{10}	95% ICB
EO1 – EO2	0,51	0,42	$2,66 \times 10^{10}$	[0,39, 0,61]
Ansiedad	0,39	0,28	126527,92	[0,26, 0,51]
Motivación	0,25	0,13	20,04	[0,10, 0,38]
Agrado	0,56	0,46	$2,28 \times 10^{12}$	[0,44, 0,65]
Seguridad	0,46	0,35	$3,46 \times 10^7$	[0,33, 0,57]
Utilidad	0,47	0,36	$7,80 \times 10^7$	[0,34, 0,57]
ER1 – ER2	0,49	0,39	$2,55 \times 10^9$	[0,37, 0,60]
Ansiedad	0,43	0,33	$3,45 \times 10^6$	[0,30, 0,55]
Motivación	0,32	0,20	786,85	[0,18, 0,45]
Agrado	0,49	0,39	$1,29 \times 10^9$	[0,37, 0,60]
Seguridad	0,39	0,28	77558,75	[0,25, 0,51]
Utilidad	0,39	0,28	33134,76	[0,24, 0,50]

Nota. El límite superior para el intervalo de confianza clásico al 95% de la correlación es igual a 1 en todos los casos. El valor de p para el contraste unilateral de la correlación es inferior a 0,001 en todos los casos. EO1: primera administración de la escala original, EO2: segunda administración de la escala original, ER1: primera administración de la escala reducida, ER2: segunda administración de la escala reducida, 95% IC: límite inferior del intervalo de confianza clásico al 95% de seguridad, FB_{10} : Factor de Bayes en favor de la hipótesis alternativa de correlación positiva, 95% ICB: intervalo de credibilidad bayesiano para la correlación al 95% de seguridad.

Los resultados obtenidos muestran que las puntuaciones totales generadas por la versión original y reducida covarían tanto que más del 96% de la variabilidad en la puntuación de una de las versiones puede explicarse a partir de la otra. En la Tabla 5 se puede notar que la correlación entre una y otra puntuación se aproxima considerablemente a una correlación perfecta. Además, todos los estadísticos de correlación obtenidos, tanto en la primera como en la segunda fase del estudio, son superiores a 0,9. Es destacable que en los dos momentos temporales en los que se administró la escala la puntuación en la dimensión Ansiedad fue la que más correlacionó entre una y otra versión. Por otro lado, la puntuación de la dimensión Motivación fue la que obtuvo coeficientes de correlación menores entre cada versión de la escala.

Tabla 5
Correlaciones entre las Dimensiones de la Escala Original y Reducida en las Dos Administraciones

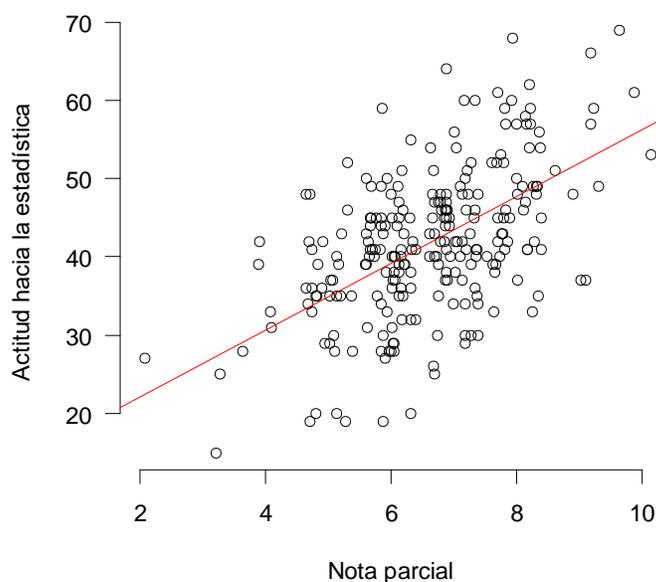
Administración y dimensión		<i>R</i>	95% IC	FB_{10}	95% ICB
O – R		0,98	0,98	$1,48 \times 10^{122}$	[0,98, 0,99]
Fase 1	Ansiedad O-R	0,97	0,96	$1,12 \times 10^{101}$	[0,96, 0,98]
	Motivación O-R	0,90	0,87	$6,94 \times 10^{58}$	[0,86, 0,92]
	Agrado O-R	0,94	0,93	$2,03 \times 10^{79}$	[0,92, 0,96]
	Seguridad O-R	0,95	0,94	$2,91 \times 10^{85}$	[0,94, 0,96]
	Utilidad O-R	0,91	0,88	$1,36 \times 10^{61}$	[0,87, 0,93]
Fase 2	Ansiedad O-R	0,97	0,96	$2,11 \times 10^{99}$	[0,96, 0,98]
	Motivación O-R	0,91	0,88	$2,12 \times 10^{61}$	[0,87, 0,93]
	Agrado O-R	0,95	0,93	$1,14 \times 10^{80}$	[0,93, 0,96]
	Seguridad O-R	0,96	0,95	$5,76 \times 10^{88}$	[0,94, 0,97]
	Utilidad O-R	0,95	0,93	$2,83 \times 10^{81}$	[0,93, 0,96]

Nota. El límite superior para el intervalo de confianza clásico al 95% de la correlación es igual a 1 en todos los casos. El valor de *p* para el contraste unilateral de la correlación es inferior a 0,001 en todos los casos. O: escala original, R: escala reducida, 95% IC: límite inferior del intervalo de confianza clásico al 95% de seguridad, FB_{10} : Factor de Bayes en favor de la hipótesis alternativa de correlación positiva, 95% ICB: intervalo de credibilidad bayesiano para la correlación al 95% de seguridad.

Como se puede apreciar en la Figura 2, a medida que aumenta la nota parcial de la asignatura de análisis de datos (cuando se había superado el 50% de los contenidos de esta) mayor es la actitud favorable hacia la estadística, $r = 0,56$, $p < 0,001$, $FB_{10} = 1,28 \times 10^{19}$. También se observó que a mayor nota en el examen parcial mayor puntuación en la dimensión Motivación, $r = 0,27$, $p < 0,001$, $FB_{10} = 556,19$, mayor agrado promedio hacia la estadística, $r = 0,36$, $p < 0,001$, $FB_{10} = 1,05 \times 10^6$, mayor seguridad percibida en los contenidos de estadística, $r = 0,54$, $p < 0,001$, $FB_{10} = 1,0 \times 10^{17}$, mayor valor positivo atribuido a la estadística, ($r = 0,26$, $p < 0,001$, $FB_{10} = 360.05$), y menor nivel de ansiedad hacia los contenidos de la estadística, $r = -0,52$, $p < 0,001$, $FB_{10} = 3,50 \times 10^{15}$. También se observa en la Tabla 6 que los niveles de actitud hacia la estadística mejoran en la medida en que se obtienen resultados más positivos en la evaluación final de la asignatura, $F(4, 377) = 30,07$, $p < 0,001$, $\eta^2 = 0,24$, $FB_M = 2,38 \times 10^{18}$. Como se puede notar en la Tabla 6, el patrón se repite para las dimensiones Ansiedad, $F(4, 377) = 27,41$, $p < 0,001$, $\eta^2 = 0,23$, $FB_M = 10,0 \times 10^{16}$, Motivación, $F(4, 377) = 6,62$, $p < 0,001$, $\eta^2 = 0,07$, $FB_M = 643,79$, Agrado, $F(4, 377) = 14,22$, $p < 0,001$, $\eta^2 = 0,13$, $FB_M = 5,25 \times 10^7$, Seguridad, $F(4, 377) = 27,45$, $p < 0,001$, $\eta^2 = 0,23$, $FB_M = 4,53 \times 10^{16}$, y Valor, $F(4, 377) = 20,71$, $p = 0,006$, $\eta^2 = 0,04$, $FB_M = 3,46$.

Figura 2

Relación entre la Nota Parcial de la Asignatura y la Puntuación en la Versión Reducida de la Escala de Actitud hacia la Estadística de Auzmendi (1992)



Nota. La línea roja representa la línea del modelo de regresión lineal que predice la actitud hacia la estadística en función de la nota en el examen parcial.

Tabla 6

Puntuaciones Medias y Desviaciones Típicas para la Puntuación de la Versión Reducida de la Escala de Actitudes hacia la Estadística de Auzmendi (1992) en Función de la Calificación Final en la Asignatura

Escala total y dimensión	Nota final en la asignatura				
	Suspenseo	Aprobado	Notable	Sobresaliente	Matrícula de honor
Actitud total	31,2 (7,3) [26,6, 35,9]	39,3 (8,1) [38,1, 40,5]	46,7 (8,9) [45,4, 48,0]	51,7 (9,7) [47,2, 56,2]	59,2 (8,7) [50,9, 67,6]
Ansiedad	4,9 (1,7) [3,3, 6,5]	7,7 (2,8) [7,3, 8,1]	10,4 (3,2) [9,9, 10,8]	11,6 (3,1) [10,0, 13,1]	12,5 (3,1) [9,6, 15,4]
Motivación	7,8 (2,7) [6,5, 9,0]	8,1 (2,4) [7,8, 8,4]	9,1 (2,2) [8,7, 9,4]	10,0 (2,1) [8,8, 11,2]	11,0 (2,2) [8,7, 13,3]
Agrado	4,0 (1,6) [2,7, 5,3]	5,5 (2,2) [5,2, 5,9]	6,9 (2,6) [6,5, 7,3]	7,8 (2,7) [6,5, 9,0]	10,2 (3,0) [7,9, 12,6]
Seguridad	6,3 (2,4) [5,0, 7,6]	8,4 (2,2) [8,1, 8,7]	10,2 (2,5) [9,8, 10,6]	11,9 (2,5) [10,7, 13,2]	13,7 (1,5) [11,5, 16,0]
Valor	8,2 (2,6) [6,9, 9,5]	9,5 (2,5) [9,2, 9,9]	10,1 (2,2) [9,8, 10,5]	10,4 (2,2) [9,2, 11,7]	11,7 (2,1) [9,4, 14,1]

Nota. La desviación típica aparece entre paréntesis y el intervalo de confianza para la media al 95% aparece entre corchetes. La puntuación en la dimensión Ansiedad se presenta invertida y, por tanto, un mayor valor representa menor nivel de ansiedad.

Discusión

Los resultados muestran que la versión reducida de la escala de medición de las actitudes hacia la estadística es un instrumento que goza de propiedades psicométricas comparables, e incluso mejores en algunos aspectos, que las de la escala original (Auzmendi, 1992). La reducción de ítems sugerida por Ruiz-

Ruano García, Puga et al. (2018), con base en un estudio cualitativo de la escala y que se ha analizado en este trabajo, produce una escala tan buena como la original e, incluso, más parsimoniosa desde el punto de vista estadístico, que supone una mínima pérdida de información. Por ejemplo, si se atiende al coeficiente de determinación, más del 96% de la variabilidad de la escala original puede ser explicada por la variabilidad en la versión reducida. La escala fue desarrollada en el contexto educativo universitario y, por tanto, puede ser un instrumento apropiado para evaluar la actitud hacia la estadística en la universidad o en cursos preuniversitarios. La reducción del número de ítems, por tanto, hace de esta escala un instrumento que permitiría reducir tiempos y costes de administración en contextos educativos con propósitos diagnósticos o de optimización del proceso de enseñanza-aprendizaje.

Como han puesto de manifiesto los resultados, la consistencia interna es comparable con la de la escala original. La reducción observada en los coeficientes de consistencia interna puede explicarse por la reducción en el número de ítems, ya que, por ejemplo, el coeficiente alfa de Cronbach está en función del número de ítems de la escala o subescala (Barrios & Cosculluela, 2013). La estabilidad temporal de la medida de actitudes hacia la estadística por medio de la escala reducida puede considerarse mejor que la de la escala original. Aunque las correlaciones entre las puntuaciones de la primera y segunda administración son todas positivas, son de magnitud moderada. Es destacable la mayor magnitud de la correlación entre la primera y segunda medición en la dimensión Motivación en la versión reducida, así como el factor de Bayes que se le asocia y que sugiere una asociación genuina entre ambas puntuaciones. En cualquier caso, estos estadísticos moderados de estabilidad temporal pueden explicarse por el cambio en el constructo que podrían haber experimentado los alumnos a lo largo del desarrollo de la asignatura. Estas evidencias sobre la estabilidad temporal de la escala son particularmente destacables, ya que aportan un valor añadido que algunas veces no son aportadas en otros estudios. Por otro lado, las relaciones observadas entre las puntuaciones de la versión reducida de la escala y las notas obtenidas en la asignatura de análisis de datos pueden considerarse como evidencias en favor de la validez de constructo. Tanto la nota del examen parcial, que se realizó cuando se había tratado el 50% de los contenidos de la asignatura, como la nota final obtenida en la asignatura se relacionan con la puntuación global de la escala y con sus dimensiones en el sentido esperado.

Los resultados de este trabajo son consistentes con estudios previos en los que se ha observado que la dimensión Ansiedad es la más consistente y fiable de la escala (véase, por ejemplo, Darias Morales, 2000; Ruiz-Ruano García et al., 2020; Vilà Baños & Rubio Hurtado, 2016). En este estudio se observó que esta dimensión de la escala reducida es la que presenta mayor consistencia interna y es la más congruente con la versión original de la escala. Por tanto, a la espera de estudios posteriores que corroboren estos resultados, este subcomponente de la escala podría considerarse apropiado para evaluar los niveles de ansiedad hacia la estadística en contextos educativos de manera precisa (Chew et al., 2017). Esto permitiría la valoración del grado en que diferentes propuestas docentes o enfoques de aprendizaje tienden a reducir la ansiedad hacia la estadística y optimizan el proceso de enseñanza-aprendizaje.

Uno de los puntos débiles de la escala reducida, como se comentó anteriormente, es la modesta estabilidad temporal de la medida. Este resultado podría ser explicado por el tiempo tan grande (todo un cuatrimestre) que transcurrió entre la primera y la segunda administración. Futuros estudios deberían explorar este aspecto de la fiabilidad y contrastar si la versión reducida de la escala tiene un buen comportamiento en otros contextos educativos (otras titulaciones o niveles, por ejemplo). Como ha sido puesto de manifiesto por los análisis paralelos, la dimensionalidad de la versión reducida es susceptible de ser estudiada en mayor profundidad. Por ello, estudios futuros con muestras más amplias y representativas tendrían que abordar esta cuestión y valorar el grado en que la estructura factorial es congruente con la de la escala original. Otra limitación de este estudio tiene que ver con el tipo de muestreo llevado a cabo. El muestreo no-probabilístico llevado a cabo para obtener la muestra de participantes para este estudio implica que los resultados no son del todo extrapolables a otras universidades o a otras regiones del mundo. Por tanto, sería deseable que se llevaran a cabo investigaciones orientadas a contrastar los resultados aquí presentados en diferentes contextos socioculturales.

De manera general, los resultados sugieren que la versión reducida de la escala para medir actitudes hacia la estadística tiene un buen comportamiento psicométrico y es más parsimoniosa que la versión original (Auzmendi, 1992). Como consecuencia, esta escala podría utilizarse para evaluar u optimizar el proceso de enseñanza-aprendizaje de la estadística. Del mismo modo, esta escala reducida también podría ser utilizada para identificar los factores de riesgo que se asocian con actitudes negativas y minimizarlos en pos de un

aprendizaje saludable de la estadística. En definitiva, como sugiere Gigerenzer (1998, 2004), el objetivo a perseguir sería conseguir una alfabetización estadística global en la que se razonase sensatamente sobre los datos estadísticos, evitando un uso ritual y "descerebrado" de las técnicas o herramientas estadísticas. Aunque, como han señalado algunos autores (Estrada et al., 2013), la actitud negativa que muchos estudiantes tienen hacia la estadística proviene de su naturaleza u origen matemático, debería trabajarse para "alejar" parcialmente la estadística de la matemática pura. Si se hace esto, como sugieren Tintle et al. (2015), se podría conseguir que los alumnos se aproximen mucho más y mejor al pensamiento estadístico.

Referencias

- Acón Araya, S. & Salazar Ramos, M. (2020). Actitudes de universitarios hacia la estadística como materia de estudio y herramienta para analizar datos. *Revista Interamericana de Psicología*, 54(1), Artículo e1140.
- Aczel, B., Hoekstra R., Gelman, A., Wagenmakers, E. -J., Klugkist, I. G., Rouder, J. N., Vandekerckhove, J., Lee, M. D., Morey, M. D., Vanpaemel, W., Dienes, Z., & van Ravenzwaaij, D. (2020). Discussion points for Bayesian inference. *Nature Human Behaviour*, 4(6), 561-563. <https://doi.org/10.1038/s41562-019-0807-z>
- Ajzen, I. & Fishbein, M. (2005). The influence of attitudes on behavior. En D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173-221). Lawrence Erlbaum.
- Allport, G. W. (1935). Attitudes. En C. Murchison (Ed.), *Handbook of social psychology* (pp. 798-844). Clark University Press.
- Altman, N. & Krzywinski, M. (2017). P values and the search for significance. *Nature Methods*, 14(1), 3-4. <https://doi.org/10.1038/nmeth.4120>
- Auzmendi, E. (1992). *Las actitudes hacia la matemática-estadística en las enseñanzas medias y universitarias. Características y medición*. Mensajero.
- Bakan, D. (1966). The test of significance in psychological research. *Psychological Bulletin*, 66(6), 423-437. <https://doi.org/10.1037/h0020412>
- Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, 533, 452-454. <https://doi.org/10.1038/533452a>
- Barrios, M. & Coscolluela, A. (2013). Fiabilidad. En J. Meneses (Coord.), *Psicometría* (pp. 75-140). Editorial Universitat Oberta de Catalunya.
- Barron, K. E. & Apple, K. J. (2014). Debating curricular strategies for teaching statistics and research methods: What does the current evidence suggest? *Teaching of Psychology*, 41(3), 187-194. <https://doi.org/10.1177/0098628314537967>
- Bates, J. A. (1991). Teaching hypothesis testing by debunking a demonstration of telepathy. *Teaching of Psychology*, 18(2), 94-97. https://doi.org/10.1207/s15328023top1802_6
- Bishop, D. (2020). How scientists can stop fooling themselves over statistics. *Nature*, 584, 9. <https://doi.org/10.1038/d41586-020-02275-8>
- Bologna, E. L. & Vaiman, M. (2013). Actitudes, experiencia previa y nivel de logro en estadística en la carrera de psicología. En J. M. Contreras, G. R. Cañadas, M. M. Gea & P. Arteaga (Eds.), *Actas de las Primeras Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria* (pp. 91-103). Universidad de Granada, Facultad de Ciencias de la Educación.
- Bond, M. (2009). Decision-making: Risk school. *Nature*, 461, 1189-1192. <https://doi.org/10.1038/4611189a>
- Chew, P. K. H., Swinbourne, A. & Dillon, D. B. (2017). An absence of attentional bias: Statistics anxiety is unique among anxieties. *Journal of Articles in Support of the Null Hypothesis*, 13(2), 91-112. <https://www.jasnh.com/pdf/Vol13-No2-article6.pdf>
- Darias Morales, E. J. (2000). Escala de actitudes hacia la estadística. *Psicothema*, 12(Supl. 2), 175-178. <https://www.psicothema.com/pdf/542.pdf>
- Dunn, T. J., Baguley, T. & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105, 399-412. <https://doi.org/10.1111/bjop.12046>
- Estrada, A. & Batanero, C. (2015). Construcción de una escala de actitudes hacia la probabilidad y su enseñanza para profesores. En C. Fernández Verdú, M. Molina González & N. Planas Raig (Eds.), *Investigación en educación matemática XIX* (pp. 239-247). Sociedad Española de Investigación en Educación Matemática. http://rua.ua.es/dspace/bitstream/10045/51393/1/2015-Actas-XIX-SEIEM_22.pdf
- Estrada, A., Batanero, C. & Fortuny, J. M. (2004). Un estudio comparado de las actitudes hacia la estadística en profesores en formación y en ejercicio. *Enseñanza de las Ciencias*, 22(2), 263-273.
- Estrada, A., Bazán, J. & Aparicio, A. (2013). Evaluación de las propiedades psicométricas de una escala de actitudes hacia la estadística en profesores. *AIEM. Avances en Investigación en Educación Matemática*, 3, 5-23.
- Falissard, B. (2011). *Analysis of questionnaire data with R*. Chapman & Hall. <https://doi.org/10.1201/b11190>
- Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Addison-Wesley.
- Gao, J. (2020). P-values — A chronic conundrum. *BMC Medical Research Methodology*, 20(167), 1-8. <https://doi.org/10.1186/s12874-020-01051-6>
- Gigerenzer, G. (1998). We need statistical thinking, not statistical rituals. *Behavioral and Brain Sciences*, 21(2), 199-200. <https://doi.org/10.1017/S0140525X98281167>
- Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, 33 (5), 587-606. <https://doi.org/10.1016/j.soec.2004.09.033>
- Gorvine, B. J. & Smith, H. D. (2015). Predicting student success in a psychological statistics course emphasizing collaborative learning. *Teaching of Psychology*, 42(1), 56-59. <https://doi.org/10.1177/0098628314562679>
- Haller, H. & Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research Online*, 7(1), 1-20. https://www.krigolonteaching.com/uploads/4/3/8/4/43848243/haller_krauss_2002.pdf
- Hartnett, J. L. (2013). Stats on the cheap: Using free and inexpensive Internet resources to enhance the teaching of statistics and research methods. *Teaching of Psychology*, 40(1), 52-55. <https://doi.org/10.1177/0098628312465865>
- Hartwright, C. E., Looi, C. Y., Sella, F., Inuggi, A., Santos, F. H., González-Salinas, C., García Santos, J. M., Kadosh, R. C. & Fuentes, L. J. (2018). The neurocognitive architecture of individual differences in math anxiety in typical children. *Scientific Reports*, 8, Artículo 8500. <https://doi.org/10.1038/s41598-018-26912-5>
- JASP Team. (2020). JASP (Version 0.13.1) [Software computacional]. <https://jasp-stats.org>
- Jeffreys, H. (1948). *Theory of probability* (2ª ed.). Oxford University Press.
- Jeon, M. & De Boeck, P. (2017). Decision qualities of Bayes factor and p value-based hypothesis testing. *Psychological Methods*, 22(2), 340-360. <https://doi.org/10.1037/met0000140>

- Karpiak, C. P. (2011). Assessment of problem-based learning in the undergraduate statistics course. *Teaching of Psychology*, 38(4), 251-254. <https://doi.org/10.1177/0098628311421322>
- Kass, R. E. & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773-795. <https://doi.org/10.2307/2291091>
- Kelley, K. & Cheng, Y. (2012). Estimation of and confidence interval formation for reliability coefficients of homogeneous measurement instruments. *Methodology*, 8(2), 39-50. <https://doi.org/10.1027/1614-2241/a000036>
- Kruschke, J. K. (2021). Bayesian analysis reporting guidelines. *Nature Human Behavior*, 5, 1282-1291. <https://doi.org/10.1038/s41562-021-01177-7>
- León, O. G. (1996). Cómo entusiasmar a 100 estudiantes en la primera clase de metodología e introducir al mismo tiempo 22 conceptos fundamentales de la materia. *Psicothema*, 8(1), 221-226. <https://www.psicothema.com/pdf/18.pdf>
- López Puga, J. (2014). Using playing cards to differentiate probability interpretations. *Teaching Statistics*, 36(3), 76-78. <https://doi.org/10.1111/test.12055>
- Morales, J. F., Reboloso, E. & Moya, M. (1994). Actitudes. En J. F. Morales (Ed.), *Psicología social* (pp. 495-524). McGraw-Hill.
- Nolan, M. M., Beran, T. & Hecker, K. G. (2012). Surveys assessing students' attitudes toward statistics: A systematic review of validity and reliability. *Statistics Education Research Journal*, 11(2), 103-123.
- Nuzzo, R. (2014, February 12). Statistical errors: P values, the 'gold standard' of statistical validity, are not as reliable as many scientists assume. *Nature*, 506, 150-152. <https://doi.org/10.1038/506150a>
- Orlitzky, M. (2012). How can significance tests be deinstitutionalized? *Organizational Research Methods*, 15(2), 199-228. <https://doi.org/10.1177/1094428111428356>
- Pascual Llobell, J., Frías Navarro, M. D. & García Pérez, J. F. (2000). El procedimiento de significación estadística (NHST): su trayectoria y actualidad. *Revista de Historia de la Psicología*, 21(1), 9-26. <http://www.revistahistoriapsicologia.es/app/download/5833388711/01.+PASCUAL+LLOBELL%2C+J..pdf?t=1361574744>
- Ponsoda, V. (1990). Un punto de vista sobre la docencia de la estadística en psicología. *Estadística Española*, 31(122), 481-489.
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org>
- Rodríguez Parrondo, J. M. (2019). *Teoría de la información. Las matemáticas de la era digital*. EMSE EDAPP.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36. <https://doi.org/10.18637/jss.v048.i02>
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225-237. <https://doi.org/10.3758/PBR.16.2.225>
- Ruiz, M. A., Pardo, A. & San Martín, R. (2010). Modelos de ecuaciones estructurales. *Papeles del Psicólogo*, 31(1), 34-45. <https://www.papelesdelpsicologo.es/pdf/1794.pdf>
- Ruiz de Miguel, C. (2015). Actitudes hacia la estadística de los alumnos del grado en pedagogía, educación social, y maestro de educación infantil y maestro de educación primaria de la UCM. *Educación XXI*, 18(2), 351-374. <https://doi.org/10.5944/educxx1.12158>
- Ruiz-Ruano García, A. M., Miles, A., Marmolejo-Ramos, F., Gabriel, F. & Puga, J. L. (2020). Network analysis of attitudes towards statistics: The Canada and Spain cases. En L. Gómez Chova, A. López Martínez & C. Torre (Eds.), *INTED 2020: 14th International Technology, Education and Development Conference Proceedings* (pp. 4322-4328). International Academy of Technology, Education and Development. <https://doi.org/10.21125/inted.2020.1199>
- Ruiz-Ruano García, A. M., Palazón Pérez de los Cobos, A. & Puga, J. L. (2018). Six manipulative tasks to improve attitudes towards statistics at university. En L. Gómez Chova, A. López Martínez & C. Torre (Eds.), *ICERI 2018 Proceedings: 11th Annual International Conference of Education, Research and Innovation* (pp. 8502-8508), International Academy of Technology, Education and Development. <https://dx.doi.org/10.21125/iceri.2018.0555>
- Ruiz-Ruano García, A. M. & Puga, J. L. (2020). Cómo mejorar la comunicación de estadísticos inferenciales en ciencias de la salud. *Revista Española de Comunicación en Salud*, 11, 139-145. <https://doi.org/10.20318/recs.2020.5173>
- Ruiz-Ruano García, A. M., Puga, J. L. & Auzmendi Escribano, E. (2018). English version of the Auzmendi Scale to measure attitudes towards statistics. En L. Gómez, A. López, y I. Candel, *Proceedings of ICERI2018 Conference* (pp. 8520-8525), Sevilla. <https://dx.doi.org/10.21125/iceri.2018.0559>
- Tintle, N., Chance, B., Cobb, G., Roy, S., Swanson, T. & VanderStoep, J. (2015). Combating anti-statistical thinking using simulation-based methods throughout the undergraduate curriculum. *The American Statistician*, 69(4), 362-370. <https://doi.org/10.1080/00031305.2015.1081619>
- Vilà Baños, R. & Rubio Hurtado, M. J. (2016). Actitudes hacia la estadística en el alumnado del grado de pedagogía de la Universidad de Barcelona. *Revista de Docencia Universitaria*, 14(1), 131-149. <https://doi.org/10.4995/redu.2016.5766>
- Wasserstein, R. L. & Lazar, N. A. (2016). The ASA statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2), 129-133. <https://doi.org/10.1080/00031305.2016.1154108>

Fecha de recepción: Noviembre de 2020.

Fecha de aceptación: Junio de 2022.