

UNA REVISIÓN DE LAS EVIDENCIAS DE FIABILIDAD Y VALIDEZ DE LOS CUESTIONARIOS DE ACTITUDES Y ANSIEDAD HACIA LA ESTADÍSTICA

JOSE CARMONA MÁRQUEZ
 Universidad de Huelva, España
 carmona@uhu.es

RESUMEN

La importancia del dominio afectivo en el ámbito de la educación estadística es en la actualidad ampliamente reconocida. Con objeto de evaluar la influencia de ese componente afectivo en la formación estadística de los estudiantes, se han elaborado en las dos últimas décadas un número importante de instrumentos para la medición de las actitudes y de la ansiedad hacia la estadística. Sin embargo, muchos de estos instrumentos han sido muy poco investigados y, en general, sus propiedades psicométricas no parecen estar bien establecidas. El objetivo de este estudio es realizar una revisión sistemática de los trabajos de investigación que aportan evidencias empíricas acerca de la fiabilidad y validez de dichos instrumentos. Los resultados de esta revisión muestran ciertas carencias que hacen recomendables nuevos estudios acerca de estos instrumentos de medición de las actitudes y la ansiedad hacia la estadística. Finalizamos proponiendo algunas líneas de investigación que creemos pueden tener cierto impacto en la investigación futura sobre el tema.

Palabras clave: Medición de la actitud; Ansiedad; Estadística; Revisión de la literatura

ABSTRACT

Nowadays, the importance of affectivity in statistics education is widely acknowledged. In the past two decades a large number of instruments to measure attitudes and anxiety toward statistics have been developed in order to assess the influence of emotional factor on students' training. However, many of these instruments have not been analysed in depth and their psychometric properties have not properly been assessed. The aim of this paper is revising the research works that have contributed empirical evidence for the reliability and validity of these instruments. Our results suggest the need for more research about instruments to measure attitudes and anxiety about statistics. We conclude by suggesting some priority research areas that might be relevant in future studies on this topic.

Keywords: Attitude Measurement, Anxiety, Statistics, Literature Review

1. INTRODUCCIÓN

Muchos alumnos suelen llegar a las asignaturas de estadística con preconcepciones y actitudes negativas hacia la materia, asociadas habitualmente con niveles altos de ansiedad cuando se enfrentan a las clases, ejercicios o exámenes de estas asignaturas. Por ejemplo, en un estudio sobre la prevalencia de la ansiedad hacia la estadística, Onwuegbuzie (en prensa) estima que aproximadamente un 75% de los alumnos experimentan niveles elevados de ansiedad. Otros autores han señalado que los alumnos suelen percibir estas asignaturas como obstáculos en el camino para la obtención del título (Perney y Ravid, 1990). Un indicador claro de esta situación es el apodo con el

que muchos alumnos en las universidades anglosajonas suelen referirse a la asignatura de estadística: ‘sadística’ -*sadistics*- (Rosenthal, 1992). Dillon (1982) ha llegado incluso a etiquetar como ‘estadísticofobia’ -*statisticophobia*- esos sentimientos de los alumnos.

A pesar de la importancia del dominio afectivo para un aprendizaje adecuado y un uso efectivo de la estadística, existen pocas investigaciones que aborden el estudio de los constructos de las actitudes y la ansiedad hacia la estadística. En este sentido, sigue siendo válida la afirmación de Gal y Ginsburg (1994) cuando señalaban que “el conjunto de investigaciones sobre las actitudes, las creencias y los afectos relacionados directamente con la educación estadística es reducido y problemático” (párrafo 17). Uno de los problemas de este campo de estudios es la falta de consenso acerca de cuáles son los instrumentos de medida adecuados para medir estos constructos. Como en otras áreas de investigación que están en sus etapas iniciales de desarrollo, existe cierta proliferación de instrumentos de medida que suelen ser usados en pocas ocasiones y cuyo fundamento teórico es poco claro (Schuessler, 1993).

El avance de la investigación del dominio afectivo en la educación estadística requiere, sin embargo, de instrumentos que proporcionen datos fiables y válidos acerca de las actitudes y la ansiedad hacia la estadística. En la literatura sobre educación estadística ya existen algunas revisiones de los instrumentos elaborados para medir estos constructos. Así, por ejemplo, Cashin y Elmore (1997) presentan una recopilación de la mayoría de investigaciones realizadas hasta esa fecha, tanto de actitudes como de ansiedad hacia la estadística, con especial énfasis en las evidencias de fiabilidad y validez presentadas en los mismos. El trabajo de Estrada (2001), en cambio, se centra en el análisis de los instrumentos para medir actitudes, más desde un punto de vista teórico que empírico. De la misma forma, Onwuegbuzie y Wilson (2003) presentan una revisión muy interesante del constructo ansiedad hacia la estadística, en la que, sin embargo, no se profundiza en las características psicométricas de los cuestionarios elaborados para medir dicha ansiedad.

El objetivo de este trabajo es poner a disposición de los investigadores una revisión de la información disponible acerca de los diferentes instrumentos propuestos para medir estos constructos. Se pretende actualizar el trabajo de Cashin y Elmore (1997) y complementar los trabajos de Estrada (2001) y Onwuegbuzie y Wilson (2003) con una revisión que incluya artículos de revista, comunicaciones a congresos y tesis doctorales obtenidos a través de búsquedas en las bases de datos PSYCINFO (Base de datos de información psicológica de la American Psychological Association) y ERIC (Educational Resources Information Center). Fruto de esta revisión se encontraron 112 trabajos de investigación en los que se usaba alguna medida de las actitudes o ansiedad hacia la estadística.

Esta revisión pretende ser un intento sistemático de presentar las evidencias de fiabilidad y validez de estos instrumentos que permita obtener conclusiones acerca de los defectos o virtudes de los mismos. Antes de abordar el análisis de los resultados, hemos planteado dos apartados iniciales cuyo objetivo es introducir conceptualmente las actitudes y la ansiedad hacia la estadística y describir las características principales de los instrumentos de medida analizados. A continuación se incluyen los dos apartados en los que se da cuenta de los resultados de la revisión: las evidencias de fiabilidad y de validez de las medidas de actitudes o ansiedad hacia la estadística. Por último se extraen algunas conclusiones que creemos representan un punto de partida conveniente para las próximas investigaciones en este campo.

2. LAS ACTITUDES Y LA ANSIEDAD HACIA LA ESTADÍSTICA

Tal como señalan la mayoría de los teóricos de la psicometría, la medición de un constructo requiere de una delimitación conceptual precisa del mismo. Para que se puedan elaborar adecuadamente los ítems de un instrumento de medida, es necesario especificar cuáles son las características comunes y diferenciales de esos ítems. Un paso previo a la elaboración de instrumentos de medida debe ser, por tanto, definir el constructo y especificar las dimensiones o facetas que lo conforman. Aunque el objetivo de este trabajo no es elaborar un instrumento de medida, sí lo es evaluar la calidad de instrumentos elaborados. Por ello, en lo que sigue describiremos brevemente qué vamos a entender por actitudes y ansiedad hacia la estadística.

Según Eagly y Chaiken (1998) una actitud es una “tendencia psicológica que es expresada a través de la evaluación de una entidad particular favorable o desfavorablemente en cierto grado”

(p.269). En el caso de las actitudes hacia la estadística esa tendencia se forma a lo largo del tiempo como consecuencia de las emociones y sentimientos experimentados en el contexto del aprendizaje de las matemáticas y la estadística (Gal, Ginsburg y Schau, 1997).

Esta concepción general de las actitudes hacia la estadística es compartida por la mayoría de autores, aunque no existe acuerdo acerca de cuáles son las dimensiones que estructuran este dominio psicológico. Las dos visiones más aceptadas son las propuestas por Wise (1985) y por Schau, Stevens, Dauphinee y del Vecchio (1995). Para Wise (1985) existen dos dimensiones relevantes en el dominio de las actitudes de los alumnos hacia la estadística: las actitudes hacia la asignatura de estadística en que están matriculados y las actitudes hacia el uso de la estadística en su campo de estudios. Schau et al. (1995), por otra parte, distinguen cuatro dimensiones en las actitudes hacia la estadística: (a) Afectos -sentimientos positivos o negativos en relación a la estadística-, (b) Competencia cognitiva -actitudes acerca del conocimiento y las habilidades intelectuales aplicadas a la estadística-, (c) Valor -actitudes acerca de la utilidad, relevancia y valía de la estadística en la vida personal y profesional-, y (d) Dificultad -actitudes acerca de la dificultad de la estadística como materia-. Como veremos más adelante, estas propuestas han servido como base para la elaboración de dos de los cuestionarios de actitudes hacia la estadística más usados en la actualidad.

La ansiedad hacia la estadística puede ser definida como una fobia específica, es decir, como un miedo desmesurado hacia un objeto, en este caso la estadística. Tal como señala Zeidner (1991) la ansiedad hacia la estadística puede manifestarse a través de la aparición de una preocupación excesiva, pensamientos perturbadores, tensión y cierta excitación fisiológica, en determinadas situaciones académicas. Este carácter situacional es una de las características definitorias de la ansiedad y representa un elemento diferenciador con respecto a las actitudes hacia la estadística. Además, constituye el fundamento de la dimensionalidad de este constructo, de manera que es habitual distinguir entre ansiedad hacia el examen, ansiedad en clase o ansiedad al resolver problemas, entre otros tipos de ansiedad hacia la estadística que pueden surgir en situaciones académicas específicas.

3. INSTRUMENTOS DE MEDIDA DE LAS ACTITUDES Y LA ANSIEDAD HACIA LA ESTADÍSTICA

El análisis de la literatura existente arroja un resultado de 17 instrumentos de medida de las actitudes y la ansiedad hacia la estadística. Exceptuando dos instrumentos elaborados a partir de escalas bipolares, a la manera del diferencial semántico de Osgood (Birenbaum y Eylath, 1994; Green, 1993), todos los instrumentos revisados son escalas tipo Likert. En lo que sigue vamos a describir brevemente estos cuestionarios, poniendo un mayor énfasis en aquellos que han sido usados más frecuentemente.

3.1 CUESTIONARIOS DE ACTITUDES HACIA LA ESTADÍSTICA

Los estudios acerca de las actitudes hacia la estadística se remontan al menos a mediados de los años 50 con un trabajo de Bending y Hughes (1954). Esta temática, sin embargo, recibió muy poca atención hasta los años 80 cuando varios autores diseñaron algunos de los instrumentos específicos para la medida de estas actitudes (Roberts y Bilderback, 1980; Wise, 1985). Los cuestionarios más usuales para medir las actitudes hacia la estadística son el SAS -*Statistics Attitude Survey*- de Roberts y Bilderback (1980), el ATS -*Attitudes Toward Statistics scale*- de Wise (1985) y el SATS -*Survey of Attitudes Toward Statistics*- de Schau et al. (1995).

Statistics Attitude Survey (SAS)

Roberts y Bilderback (1980) pretendían elaborar una escala afectiva que por estar “formulada en lenguaje estadístico” (p. 236) fuera más relevante que otros instrumentos de medida de tipo afectivo –escalas de actitudes hacia las matemáticas- para la predicción de las notas en las asignaturas de estadística. A partir de un conjunto inicial de 50 ítems, desarrollados tomando como referencia el dominio de contenido presentado por Dutton (1951) para su escala de actitudes hacia las matemáticas,

seleccionaron los 34 ítems que componen el SAS, de manera que tuvieran una correlación alta con el total de la escala. Cada uno de estos ítems tiene cinco opciones de respuesta, desde ‘Totalmente de acuerdo’ hasta ‘Totalmente en desacuerdo’.

A pesar de que el SAS “fue construido teniendo en cuenta varios conceptos tales como la competencia percibida en estadística y la utilidad de la estadística, no se supuso que era factorialmente complejo” (Roberts y Reese, 1987, p. 759). En general parece deducirse de sus escritos que Roberts considera que este cuestionario como esencialmente unidimensional, aunque en ocasiones use argumentos no demasiado convincentes. Así, por ejemplo, en Roberts y Bilderback (1980) se señala que las altas estimaciones de la consistencia interna indican que “los ítems miden el mismo constructo” (p. 237), a pesar de las advertencias que se pueden encontrar en la literatura psicométrica acerca de que las estimaciones de consistencia interna no son buenos índices de la dimensionalidad de una prueba (Crocker y Algina, 1986).

Según Wise (1985), la inspección de los ítems del SAS revela que al menos un tercio de ellos abordan cuestiones como la capacidad de los estudiantes en la solución de problemas de estadística o la comprensión de conceptos estadísticos. Para Wise estos ítems son inapropiados por dos razones: en primer lugar parecen estar midiendo más el logro que las actitudes de los estudiantes y, en segundo lugar, son inapropiados para alumnos que acaban de comenzar la asignatura de estadística.

Attitudes Toward Statistics scale (ATS)

Los ítems del ATS difieren de los del SAS en tres aspectos (Wise, 1985): (a) se elaboraron para que sean por naturaleza actitudinales, (b) pretenden ser contestables por los estudiantes tanto en el primero como en el último día de clase, y (c) se construyeron para medir dos dominios separados, a saber, las actitudes de los estudiantes hacia la asignatura en la que están matriculados (ATS-Asignatura) y sus actitudes hacia el uso de la estadística en su campo de estudios (ATS-Campo). A partir de un conjunto inicial de 40 ítems con cinco opciones de respuesta cada uno, desde ‘Totalmente de acuerdo’ hasta ‘Totalmente en desacuerdo’, se seleccionaron los 29 ítems que componen el cuestionario definitivo, a través de análisis realizados por jueces expertos acerca la adecuación del contenido y de su aplicabilidad el primer día de clase, y de análisis empíricos teniendo en cuenta la correlación corregida ítem-total.

Survey of Attitudes Toward Statistics (SATS)

Las características que Schau et al. (1995) sugieren para un uso óptimo de un cuestionario de actitudes hacia la estadística, tanto en situaciones de investigación como de instrucción son las siguientes:

- (a) que cubran las dimensiones más importantes de las actitudes hacia la estadística; (b) que sean aplicables en la mayoría de los departamentos que ofrecen cursos de introducción a la estadística y sirvan como medidas relevantes a lo largo del curso con sólo cambios menores en el tiempo verbal; y (c) sean cortas, de modo que su administración implique un mínimo tiempo de clase e incluyan ítems que midan tanto actitudes positivas como negativas (p. 869).

Estos autores señalan, además, que tanto el desarrollo de la encuesta como la validación de contenido deben tener en cuenta a los estudiantes, puesto que son ellos los que rellenan el cuestionario. Por último, señalan que la estructura del instrumento resultante debe ser evaluada a través de técnicas de análisis confirmatorio, como el análisis factorial confirmatorio. Según Schau et al. (1995), ninguno de los instrumentos existentes hasta ese momento había sido elaborado teniendo en cuenta todas estas características. Para subsanar esta situación estos autores elaboraron el SATS atendiendo a todos esos factores.

A partir de una técnica de grupo nominal con alumnos y profesores de asignaturas de estadística, generaron inicialmente 80 ítems cuya estructura, alcanzada por consenso entre los participantes en la técnica grupal se compone de las cuatro dimensiones ya señaladas en el apartado anterior: Afecto, Competencia Cognitiva, Valor y Dificultad. De la muestra inicial de ítems, se seleccionan 28 ítems basándose en varios análisis: grado de acuerdo de los expertos acerca de la dimensión de la cual es indicador cada uno de los ítems; análisis empíricos clásicos basados en la correlación corregida ítem-total de la dimensión, correlación múltiple al cuadrado; valor del coeficiente alfa de Cronbach cuando

el ítem es eliminado de la escala; y análisis empírico del ajuste de los ítems a la estructura propuesta con técnicas de análisis factorial confirmatorio.

Otros cuestionarios de actitudes hacia la estadística

Otros cuestionarios elaborados para medir actitudes hacia la estadística y que han sido usados en un menor número de trabajos son los siguientes:

- a) *Students' Attitudes Toward Statistics* de Sutarso (1992), cuestionario compuesto de 24 ítems, que no parece diferir mucho de otros cuestionarios más usuales como el SAS y el ATS (Gal y Ginsburg (1994);
- b) *Statistics Attitude Scale* de McCall, Belli y Madjidi (1991) cuestionario compuesto por 20 ítems, que ha sido posteriormente usado en Glencross y Cherian (1992) y Cherian y Glencross (1997) con poblaciones sudafricanas;
- c) *Attitude Toward Statistics* de Miller, Behrens, Green y Newman (1993), instrumento compuesto de 25 ítems que pretenden medir los constructos valor de la estadística, orientación a objetivos y habilidad percibida en relación a la estadística;
- d) *Quantitative Attitudes Questionnaire* de Chang (1996), cuestionario diseñado para medir actitudes hacia la metodología cuantitativa, incluyendo la estadística pero no limitándose a ella. En su versión corta está compuesto por 20 ítems elaborados para medir cuatro dimensiones: utilidad de la metodología cuantitativa, valor de esa metodología para la investigación en ciencias sociales, eficacia o percepción de las propias habilidades en relación con lo cuantitativo, y conocimiento acerca de estos temas;
- e) *Escala de actitudes em relação à Estatística* de Cazorla, Silva, Vendramini y Brito (1999), escala unidimensional compuesta por 20 ítems y adaptada de la escala de actitudes hacia las matemáticas de Aiken y Dreger (1961).

De todos los instrumentos analizados, los únicos que, según nuestra revisión, han sido traducidos y usados en español son el SAS de Roberts y Bilderback, usado por Auzmendi (1991), Cuesta, Rifá y Herrero (2001) y Mejía (1995), el ATS de Wise, usado por Gil (1999) y el SATS de Schau et al., usado por Carmona y Moreno (1999), Estrada (2002) y Huedo, López, Martínez y Nortes (2003). Por otra parte, existen dos instrumentos elaborados directamente en español: la *EAE -Escala de Actitudes hacia la Estadística-* de Auzmendi (1991) y otro cuestionario desarrollado más recientemente con el mismo nombre por Velandrino y Parodi (1999).

La *EAE* de Auzmendi (1991) es una escala tipo Likert compuesta por 25 ítems, que ha sido elaborada para medir indistintamente actitudes hacia la estadística y hacia las matemáticas, con pequeñas modificaciones de cada ítem. Para la elaboración del cuestionario se partió de la selección de las cinco dimensiones más comunes en las escalas de actitudes hacia las matemáticas. Se seleccionaron a continuación ocho ítems por cada una de estas dimensiones y se probaron los 40 ítems resultantes en una muestra de alumnos universitarios. Se eligieron los 25 ítems definitivos, cinco por cada factor, usando como criterio de selección de ítems la saturación de éstos en el factor correspondiente en un análisis factorial. Según Auzmendi (1992) las dimensiones o factores de los que consta la *EAE* son: (a) Utilidad subjetiva que tiene para el estudiante el conocimiento de estadística; (b) Ansiedad o temor que se manifiesta ante la materia; (c) Confianza o seguridad que se tiene al enfrentarse a la estadística; (d) Agrado o disfrute que provoca el trabajo estadístico; y (e) Motivación que siente el estudiante hacia el estudio y uso de la estadística.

La *Escala de Actitudes hacia la Estadística* de Velandrino y Parodi (1999) es una escala tipo Likert desarrollada a partir de un conjunto inicial de 60 ítems. A partir de una evaluación de la adecuación del contenido por parte de jueces expertos -profesores de metodología- y de análisis clásicos de ítems se seleccionaron los ítems que componen el cuestionario definitivo. Según los autores, los resultados de un análisis factorial de la matriz de correlaciones entre ítems sirvió de apoyo a la estructura de tres dimensiones propuesta: *utilidad* percibida, *conceptualización* general de la estadística, y *disposición* y capacidad para la *formación* y preparación en estadística.

3.2 CUESTIONARIOS DE ANSIEDAD HACIA LA ESTADÍSTICA

Además de estos instrumentos para la medida de las actitudes, en la literatura se han planteado diversos instrumentos para medir la ansiedad hacia la estadística. Debemos señalar, sin embargo, que en muchos de los trabajos analizados se plantean indistintamente ambos tipos de instrumentos. A nuestro entender existen razones, tanto prácticas como teóricas, para este tratamiento indiferenciado. A un nivel práctico, la mayoría de los cuestionarios propuestos no difieren en gran medida, estando compuestos de ítems muy parecidos. Teóricamente, las actitudes suelen ser entendidas como reacciones evaluativas ante un objeto cuyos antecedentes o manifestaciones pueden ser emociones, cogniciones o conductas manifiestas. La ansiedad hacia un objeto es una reacción fundamentalmente emocional de miedo o aprensión hacia ese objeto. En este sentido, la ansiedad hacia la estadística puede ser entendida como la manifestación de una actitud negativa hacia la estadística o como un antecedente de dicha actitud. Los indicadores de ambos constructos, por tanto, pueden ser muy similares.

Una característica que introduce cierta diferenciación entre los instrumentos de actitudes y los instrumentos de ansiedad hacia la estadística es el hecho de que la ansiedad suele ser concebida como una reacción específica a la situación (Onwuegbuzie, 2000a; Zeidner, 1991), con lo que en los instrumentos para medir ansiedad suelen estar más centrados que los instrumentos de actitudes en situaciones concretas, como el examen y las clases de estadística. Los cuestionarios de ansiedad hacia la estadística más usuales son el STARS -*Statistical Anxiety Rating Scale*- de Cruise, Cash y Bolton (1985) y las diversas adaptaciones del MARS -*Mathematics Anxiety Rating Scale*- de Suinn, Edie, Nicoletti y Spinelli (1972).

Statistical Anxiety Rating Scale (STARS)

El STARS es un cuestionario compuesto de 51 ítems que evalúan la ansiedad que surge en los estudiantes en distintas situaciones académicas relacionadas con la estadística. Cada uno de los ítems consta de cinco opciones de respuesta, desde 1 ‘poca ansiedad’ hasta 5 ‘muchísima ansiedad’ provocada por la situación descrita. Cruise et al. (1985) identifican seis factores, usando un análisis factorial: (a) Valor, o percepción de los estudiantes acerca de la relevancia de la estadística; (b) Ansiedad surgida en situaciones en las que hay que interpretar datos estadísticos; (c) Ansiedad en clase; (d) Autoconcepto de cálculo, que surge cuando se tiene que resolver un problema estadístico que implica cálculos matemáticos; (e) Miedo a pedir ayuda a profesores o compañeros en diversas situaciones; y (f) Miedo a los profesores de estadística, el cual está relacionado con la percepción que se tiene del profesor.

Adaptaciones del Mathematics Anxiety Rating Scale (MARS)

En los trabajos en los que se estudia la ansiedad hacia la estadística también se han usado con bastante frecuencia algunas de las adaptaciones del MARS de Suinn et al. (1972), como el *Statistics Anxiety Inventory* de Zeidner (1991), la *Statistics Anxiety Scale* propuesta por Bessant (1997), o la versión reducida del MARS diseñada por Plake y Parker (1982) para medir ansiedad hacia las matemáticas y la estadística. Estas distintas versiones varían en función del número de ítems: así, mientras la propuesta por Plake y Parker tiene 24 ítems, el cuestionario de Zeidner tiene 40 y el de Bessant 42 ítems. En todos los casos, los ítems de los cuestionarios contienen pequeñas descripciones de situaciones, fundamentalmente académicas, relacionadas con la estadística o las matemáticas ante las cuales los sujetos deben responder en una escala de cinco puntos desde 1 ‘No me provoca ansiedad’ hasta 5 ‘Me provoca mucha ansiedad’. Los ítems de estos cuestionarios suelen ser idénticos a los del cuestionario original, con la modificación de incluir la palabra ‘estadística’ en lugar o junto a la palabra ‘matemáticas’. Todos estos cuestionarios incluyen dos dimensiones o factores denominados respectivamente ‘Ansiedad ante el aprendizaje o contenido de la estadística’ y ‘Ansiedad ante los exámenes de estadística’.

Otros cuestionarios de ansiedad hacia la estadística

Un cuestionario de ansiedad hacia la estadística con un uso menor que los anteriores es el *Statistics Anxiety Scale* de Pretorius y Norman (1992), elaborado a partir de la *Mathematics Anxiety Scale* de Betz (1978), y compuesto por 10 ítems idénticos a los de Betz, salvo en que se han cambiado las palabras referidas a matemáticas por palabras referidas a estadística.

Además en algunos de los trabajos sobre ansiedad hacia la estadística se han usado cuestionarios específicamente concebidos para medir la ansiedad hacia los exámenes como el TAI -*Test Anxiety Inventory*- de Spielberger (1980), usado por Bandalos, Yates y Thorndike-Christ (1995), Benson (1989), Musch y Bröder (1999) y Finney y Schraw (2003), y el STAS -*Statistics Test Anxiety Scale*- de Fitzgerald, Jurs y Hudson (1996) usado sólo por sus autores. El TAI es una escala tipo Likert compuesta por 20 ítems desarrollados para medir ansiedad hacia cualquier tipo de examen y que en las aplicaciones analizadas se ha usado sin modificar, con la única instrucción de que los alumnos respondan pensando en el examen de estadística. Según sus autores, el TAI mide dos dimensiones: a) Preocupación o componente cognitivo de la ansiedad hacia los exámenes que refleja las preocupaciones con la ejecución en el examen, y b) Emotividad o medida de las reacciones afectivas y fisiológicas como el nerviosismo o la tensión ante las evaluaciones. El STAS es una versión modificada del *Test Anxiety Scale* de Sarason (1972) en la que se ha reemplazado la palabra ‘examen’ por la palabra ‘estadística’; consta de 34 ítems con formato de respuesta verdadero-falso en los que se pide a los sujetos que indiquen su reacción ante diversas cuestiones relacionadas con los exámenes.

4. FIABILIDAD DE LOS INSTRUMENTOS

A continuación vamos a revisar los resultados de estos estudios desde la perspectiva de sus aportaciones al análisis de la calidad psicométrica de los cuestionarios de actitudes y ansiedad hacia la estadística. Los datos que se comentan corresponden fundamentalmente a los cinco cuestionarios más usuales -SAS, ATS, SATS, R-MARS y STARS- y a la EAE de Auzmendi, debido a que es el cuestionario en español más usado.

En casi todos los estudios revisados se informa acerca de la consistencia interna de las escalas a través del coeficiente Alfa de Cronbach. En general, estos indicadores de la consistencia interna suelen ser altos, con valores en la mayoría de los casos superiores a 0.80 -mínimo aceptable recomendado por algunos autores (Henson, 2001)- y en muchos casos superiores incluso a 0.90. Los valores mínimos y máximos del coeficiente alfa para las escalas y subescalas revisadas pueden consultarse en la Tabla 1.

A diferencia de las medidas de consistencia interna, en muy pocas ocasiones se informa acerca de la estabilidad de las puntuaciones, cuya estimación requiere administrar el instrumento en dos ocasiones, obteniéndose un coeficiente test-retest o coeficiente de correlación de Pearson entre las puntuaciones obtenidas en ambas ocasiones. Según nuestro conocimiento, sólo los trabajos de Cruise et al. (1985) y de Wise (1985) aportan medidas de la estabilidad de las puntuaciones que pueden ser interpretadas como estimaciones de la fiabilidad. Wise (1985) obtiene un coeficiente test-retest de 0.82 para la subescala Campo y de .91 para la subescala Asignatura de su cuestionario, con un tiempo entre aplicaciones de la prueba de dos semanas. Cruise et al. (1985), en cambio, dejaron transcurrir cinco semanas entre las aplicaciones de la prueba y obtuvieron unos coeficientes test-retest que varían entre 0.67 y 0.83, dependiendo de la subescala. Aunque existen otros trabajos en los que se aporta información acerca de la estabilidad de las puntuaciones (Pretorius y Norman, 1992; Shultz y Koshino, 1998), esta información no puede ser asumida a nuestro entender como un indicador de la fiabilidad debido a que el tiempo transcurrido entre aplicaciones es demasiado grande, aproximadamente tres meses en ambos casos, y a que en ese tiempo los alumnos han estado cursando una asignatura de estadística, lo cual puede haber modificado de forma natural sus actitudes.

Tabla 1. Valores mínimos y máximos del coeficiente alfa de Cronbach obtenidos en las escalas SAS, ATS, STARS, RMARS, SATS o EAE por distintos autores.

Escala	Componente	Alfa Mínimo	Alfa Máximo
SAS		0.89 (Mejía, 1995)	0.95 (Roberts y Bilderback, 1980)
ATS	Total	0.89 (Gil, 1999)	0.91 (Roberts y Reese, 1987)
	Campo	0.83 (Waters et al., 1988a)	0.94 (Perney y Ravid, 1990)
	Asignatura	0.77 (Rhoads y Hubele, 2000)	0.93 (Perney y Ravid, 1990)
STARS	Valor	0.78 (Onwuegbuzie y Daley, 1999)	0.96 (Onwuegbuzie, 2000b)
	Interpretación	0.78 (Onwuegbuzie y Daley, 1999)	0.91 (Baloglu, 2003)
	Examen	0.68 (Cruise et al., 1985)	0.91 (Baloglu, 2002)
	Cálculo	0.77 (Onwuegbuzie y Daley, 1999)	0.93 (Onwuegbuzie, 1993)
	Ayuda	0.62 (Baloglu, 2002)	0.89 (Cruise et al., 1985)
	Profesores	0.64 (Baloglu, 2003)	0.88 (Elmore et al., 1993)
Versiones de MARS	Total	0.86 (Sigurdsson, 1991)	0.98 (Plake y Parker, 1982)
	Contenido	0.94 (Zeidner, 1991)	0.94 (Zeidner, 1991)
	Examen	0.92 (Zeidner, 1991)	0.92 (Zeidner, 1991)
EAE	Total	0.89 (Sánchez-López, 1996)	0.90 (Quilter y Chester, 2001)
	Utilidad	0.64 (Auzmendi, 1991)	0.80 (Auzmendi, 1991)
	Ansiedad	0.81 (Auzmendi, 1991)	0.84 (Auzmendi, 1991)
	Seguridad	0.74 (Auzmendi, 1991)	0.84 (Auzmendi, 1991)
	Agrado	0.79 (Auzmendi, 1991)	0.83 (Auzmendi, 1991)
	Motivación	0.61 (Auzmendi, 1991)	0.71 (Auzmendi, 1991)
SATS	Afecto	0.81 (Schau et al., 1995)	0.89 (Finney y Schraw, 2003)
	Competencia	0.77 (Schau et al., 1995)	0.90 (Finney y Schraw, 2003)
	Valor	0.80 (Schau et al., 1995)	0.91 (Finney y Schraw, 2003)
	Dificultad	0.64 (Schau et al., 1995)	0.86 (Finney y Schraw, 2003)

5. VALIDEZ DE LOS INSTRUMENTOS

La validación es un proceso por el cual se aportan evidencias que apoyen la interpretación propuesta de los datos recogidos mediante la prueba, en nuestro caso del cuestionario de actitudes o ansiedad hacia la estadística. Según nuestro análisis, las evidencias aportadas en las investigaciones no están basadas en los procesos de respuesta o en las consecuencias de la evaluación, a pesar de la recomendación en este sentido de los *Standards* (AERA/APA/NCME, 1999). Una posibilidad para obtener información acerca de los procesos de respuesta es la propuesta por Gal et al. (1997), quienes consideran conveniente complementar las preguntas habituales -formato de respuesta tipo Likert- con preguntas abiertas de cuyas respuestas, para inferir las motivaciones y causas de las actitudes de los alumnos. Sin embargo, esta recomendación no ha sido seguida en la mayoría de los trabajos analizados, con la excepción de algunos trabajos recientes (Suanpang, Petocz y Kalceff, 2003; Watson, Kromrey, Ferron, Lang y Hogarty, 2003).

A continuación se presentan los resultados de las investigaciones revisadas organizados en función del tipo de evidencia que aportan en el proceso de validación de los distintos cuestionarios, que puede estar basada en (a) el contenido del cuestionario, (b) su estructura interna, o (c) su relación con otras variables.

5.1 VALIDEZ DE CONTENIDO

El análisis de aquellas investigaciones en las que se usa alguno de los seis cuestionarios revisados en mayor profundidad en este trabajo, mostró que sólo para tres de ellos se aportan evidencias de validez relativas al contenido. En estos casos, esas evidencias se obtienen en la fase de elaboración del instrumento y se presentan en los artículos originales, que describen esa fase de desarrollo de los cuestionarios. Los cuestionarios para los que se aporta este tipo de evidencias son el STARS de Cruise et al., ATS de Wise y el SATS de Schau et al.

Cruise et al. (1985) realizaron una evaluación de la adecuación del contenido de los ítems a la estructura de seis factores propuesta para el STARS presentando a cinco profesores de estadística y cinco estudiantes de doctorado descripciones de las seis subescalas e ítems potenciales para cada una de ellas. La tarea de los evaluadores era determinar a cuál de las subescalas pertenecía cada uno de los ítems. Los resultados obtenidos por estos autores mostraron unos índices de acuerdo que iban de 0.60 a 1, con una media de 0.91, lo cual fue interpretado como una evidencia de la validez aparente del contenido de los ítems.

Wise (1985) describe el proceso de evaluación del contenido de los ítems inicialmente elaborados de la siguiente manera: “*La evaluación de los 40 ítems en términos de su validez de contenido por dos profesores de introducción a la estadística en educación dio lugar a la eliminación de cinco ítems*” (p. 403). Puesto que Wise no es más explícito no podemos estar seguros acerca de qué tipo de evaluación llevan a cabo esos expertos en el contenido. Tras la eliminación de otros cinco ítems debido a sus bajos índices de discriminación, dos profesores de estadística distintos a los anteriores realizan una evaluación adicional -en este caso explícita- de la adecuación de los 30 ítems restantes para ser respondidos el primer día de clase, y cuyo resultado es la eliminación de otro ítem.

La evaluación del contenido de los ítems del SATS difiere de la realizada por Wise (1985) en dos aspectos. Por una parte, los jueces evaluadores no son sólo profesores: el grupo de jueces constaba de dos profesores de introducción a la estadística, dos alumnos graduados y dos alumnos no graduados matriculados en cursos de iniciación a la estadística. Y por otra, en Schau et al. (1995), tanto la estructura conceptual, como la relevancia de cada ítem fue determinada inductivamente a partir del consenso entre los jueces. En primer lugar, los jueces clasificaron un conjunto inicial de frases y palabras referidas a las actitudes hacia la estadística en categorías que etiquetaron y definieron con posterioridad. En esta fase “el grupo alcanzó una estructura consensuada de los ítems consistente en cuatro dimensiones” (Schau et al., 1995, p. 869), etiquetadas como Afecto, Competencia Cognitiva, Valor y Dificultad. Posteriormente esas palabras y frases fueron reescritas en 80 ítems potenciales que de nuevo fueron clasificados por cada uno de los jueces en una de las dimensiones propuestas. “Los 60 ítems (15 por categoría) que dieron lugar al mayor consenso (habitualmente 80% de acuerdo o mayor) fueron reescritos para ser usados de forma adecuada al inicio de un curso de introducción a la estadística” (Schau et al., 1995, p. 870).

Del resto de los cuestionarios descritos, sólo el *Quantitative Attitudes Questionnaire* de Chang (1996) aporta un análisis de la validez de contenido. En este caso, al igual que en el SATS, las dimensiones que conforman el constructo se derivaron a partir de las opiniones de un grupo de jueces, aunque Chang usó sólo profesores de metodología. El conjunto inicial de 65 ítems fueron revisados por esos mismos jueces para determinar su validez de contenido, reteniéndose aquellos 45 ítems que suscitaron un mayor consenso.

5.2 EVIDENCIAS BASADAS EN LA ESTRUCTURA INTERNA

El análisis de las evidencias de validez basadas en la estructura interna de una prueba consiste en el estudio del ajuste entre la estructura conceptual del constructo y las relaciones empíricas entre ítems o partes de la prueba. Aunque los análisis de datos implicados debieran depender fundamentalmente de la concepción del constructo, en las investigaciones revisadas el estudio de la estructura interna de los cuestionarios se ha realizado en todos los casos por medio de análisis factoriales de las respuestas a los ítems de esos cuestionarios, tanto exploratorios como confirmatorios.

El SAS de Roberts y Bilderback (1980) es uno de los pocos cuestionarios de actitudes o ansiedad hacia la estadística que se postulan unidimensionales, junto al *Statistics Attitude Scale* (McCall et al., 1991; Glencross y Cherian, 1992), el *Statistics Anxiety Scale* (Pretorius y Norman, 1992), y la *Escala de atitudes em relação à Estatística* (Cazorla et al., 1999). Sin embargo, cuando se ha puesto a prueba esa supuesta unidimensionalidad del SAS a través de análisis factoriales de las respuestas a sus ítems (Cuesta et al., 2001; Waters, Martelli, Zakrajsek y Popovich, 1988b) el resultado muestra como solución más satisfactoria una estructura bifactorial.

La estructura bidimensional propuesta por Wise (1985) para su ATS, sin embargo, ha sido probada en varias ocasiones usando análisis factoriales exploratorios y confirmatorios. En el trabajo original de Wise, por ejemplo, se usó un análisis factorial exploratorio con rotación varimax. La solución bifactorial obtenida dio cuenta del 49% de la varianza y los factores se interpretaron, como ya se indicó anteriormente, como actitudes hacia la asignatura y actitudes hacia el campo de la estadística. Esa estructura fue obtenida de nuevo en análisis factoriales exploratorios del ATS en los trabajos de Waters et al. (1988b) y en Woehlke (1991). En el estudio realizado por Schau, Dauphinee y Del Vecchio (1993) se aplicó un análisis factorial confirmatorio en el que se obtuvo que todos los ítems saturaban de forma significativa en su correspondiente factor, aunque el ajuste global del modelo no fue satisfactorio. Gil (1999), por otra parte, encontró resultados que no apoyaban la estructura bifactorial al realizar un análisis factorial exploratorio con rotación varimax de una versión en español del ATS, puesto que obtuvo una solución de cinco factores que explicaba aproximadamente el 51% de la varianza.

Cruise et al. (1985) realizaron un análisis factorial exploratorio con rotación varimax a partir de las respuestas de 1150 estudiantes al STARS, a partir del cual se obtuvieron precisamente los seis factores señalados en el anterior apartado de este trabajo. En un estudio más reciente, sin embargo, Baloglu (2002), usando técnicas de análisis factorial confirmatorio, concluye que el modelo de seis factores no muestra un ajuste adecuado a los datos.

Con respecto a las versiones del MARS de Suinn et al. (1972), los análisis factoriales realizados tanto en Plake y Parker (1982) como en Zeidner (1991) muestran que una estructura bifactorial da cuenta adecuadamente de las respuestas a los ítems de ambas versiones del cuestionario: las soluciones obtenidas en estos AF exploratorios dan cuenta respectivamente del 60% y del 45% de la varianza total.

La dimensionalidad del SATS de Schau et al. (1995), a diferencia del resto de cuestionarios, ha sido estudiada en distintas ocasiones a través de análisis factoriales confirmatorios. En Schau et al. (1993) se intentó confirmar la estructura de cuatro factores propuesta por los autores usando un AF confirmatorio al nivel de los ítems. Los resultados no fueron satisfactorios, debido a un pobre ajuste del modelo. Sin embargo, tanto en Schau et al. (1993) como en Schau et al. (1995) se realizó un AF confirmatorio de las puntuaciones en subgrupos de ítems *-parcels-*, formados a partir de la suma de las puntuaciones a varios ítems del mismo factor, obteniéndose en ambos casos un ajuste adecuado del modelo de cuatro factores. En Dauphinee, Schau y Stevens (1997) se encontró además que esa estructura factorial es invariante para alumnos y alumnas.

La Escala de Actitudes hacia la Estadística de Auzmendi ha sido sometida en dos ocasiones a un AF de componentes principales con rotación varimax. En la tesis doctoral de Auzmendi (1991), el resultado de dicho análisis muestra una estructura de cinco factores que da cuenta del 60.7% de la varianza total. En Sánchez-López (1996), en cambio, se obtuvo una estructura factorial distinta, constituida por cuatro factores que daban cuenta del 53.5% de la varianza.

En algunos de los estudios en los que se usan los otros cuestionarios revisados también se realizan análisis factoriales. Resumiendo, podemos señalar que mientras los resultados de los trabajos de Cazorla et al. (1999), Glencross y Cherian (1992), y Pretorius y Norman (1992) parecen apoyar la hipótesis de una estructura unifactorial de los cuestionarios respectivos; Velandrino y Parodi (1999) obtienen una estructura de tres factores para su Escala de Actitudes hacia la Estadística; Chang (1996), usando AF confirmatorio, una estructura de cuatro factores en su *Quantitative Attitudes Questionnaire*; y Miller et al. (1993) obtienen mediante análisis de componentes principales una estructura de cinco factores que muestra un ajuste aceptable a las subescalas teóricamente determinadas de su cuestionario *Attitude Toward Statistics*.

5.3 EVIDENCIAS BASADAS EN LAS RELACIONES CON OTRAS VARIABLES

Las evidencias basadas en la relación entre las puntuaciones en el test y variables externas proporcionan una de las fuentes más habituales de validación de las medidas psicológicas. La determinación de cuáles son esas variables externas depende de la teoría relativa al constructo que se

pretende medir. Aunque en nuestro caso dichas relaciones son, en principio, muy numerosas, hemos clasificado los resultados de las investigaciones revisadas en tres tipos:

1. En primer lugar, puesto que uno de los objetivos básicos de los cuestionarios de actitudes o ansiedad hacia la estadística es la predicción del rendimiento de los alumnos en las asignaturas de estadística, los intentos de validación de estos cuestionarios han incluido habitualmente el estudio de la relación entre las puntuaciones obtenidas en ellos y alguna medida del rendimiento de los alumnos, como son las notas en las asignaturas.
2. También ha sido habitual estudiar la relación que existe entre los diferentes cuestionarios de actitudes hacia la estadística. En principio es esperable que estas medidas estén altamente relacionadas, puesto que son medidas del mismo constructo.
3. Por último, existe un conjunto de investigaciones en las que se estudia la utilidad de algunas variables para predecir las actitudes hacia la estadística. En la literatura revisada hemos encontrado tres grupos de variables usadas de esta manera: características personales de los alumnos, como el género y determinadas características de personalidad; experiencia formativa en matemáticas y estadística principalmente; y pensamientos auto referidos sobre las capacidades personales.

En las investigaciones revisadas se han usado diversos métodos de análisis de datos para el estudio de esas relaciones, que van desde simples correlaciones bivariadas, hasta modelos de ecuaciones estructurales. Un modelo global de esas relaciones incluye variables predictoras de las actitudes hacia la estadística, varias medidas de las actitudes, y el rendimiento en las asignaturas de estadística, tal como se refleja en la Figura 1.

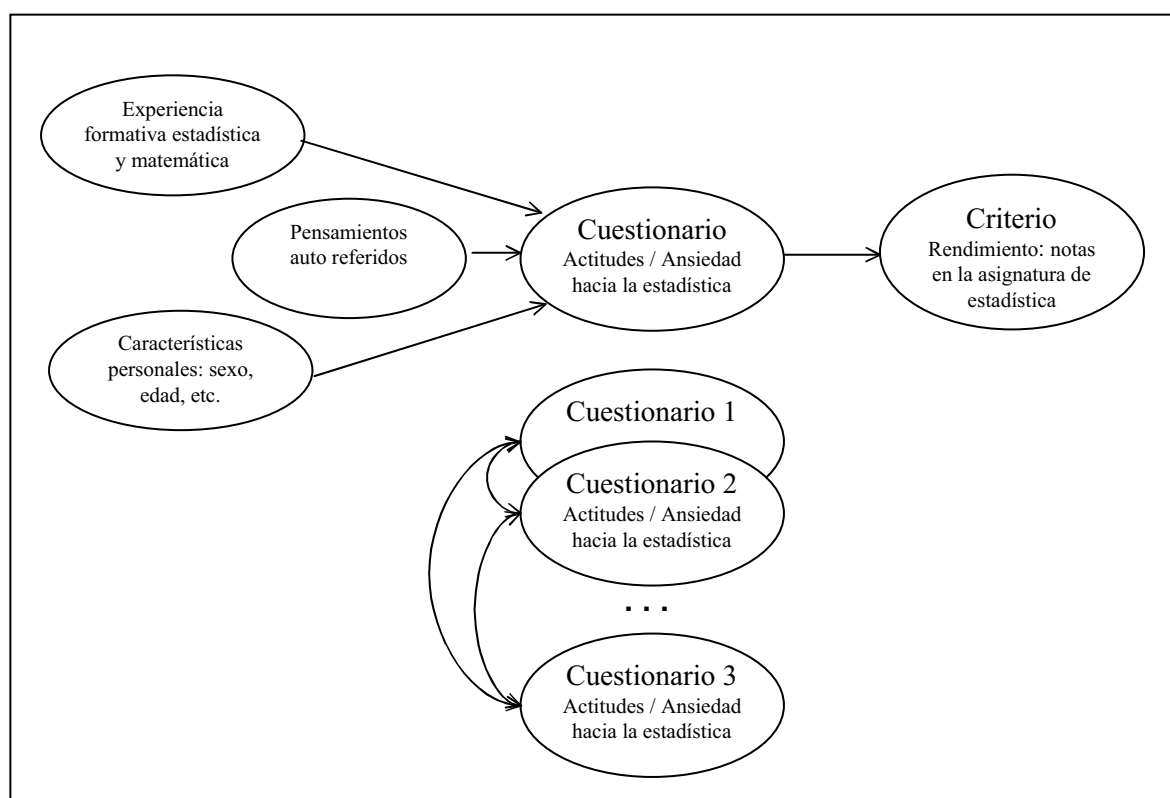


Figura 1. Modelo global de relaciones de las actitudes / ansiedad hacia la estadística con otras variables

Puesto que el objetivo de este apartado es revisar las evidencias que contribuyen a la validación de los cuestionarios de actitudes y ansiedad hacia la estadística, a continuación se presentan esas evidencias agrupadas en las tres categorías señaladas. Si atendemos únicamente al orden explicativo

planteado en el modelo global, esos tres tipos de evidencias pueden etiquetarse como evidencias de validez de pronóstico, concurrente y retrospectiva.

Evidencias de validez de pronóstico

En general, en los cuestionarios analizados se ha mostrado una relación positiva o negativa entre el rendimiento en las asignaturas de estadística y las actitudes o ansiedad hacia la estadística, que puede ser de moderada o de baja intensidad. Esto se refleja, por ejemplo, en el metaanálisis realizado por Fitzgerald (1997) en el que se obtiene una estimación global de esta relación de $r = -0.15$ a partir de 27 estudios, en los que se informa de la relación entre ansiedad hacia la estadística y rendimiento. De la misma forma, en nuestra revisión, encontramos 13 estudios en los que se informa de la relación entre el ATS de Wise y el rendimiento de los alumnos en asignaturas estadísticas, siendo la correlación promedio de 0.21. Resultados similares se obtienen con otros cuestionarios. Así, por ejemplo, Roberts y Reese (1987) encuentran una correlación entre el SAS de Roberts y Bilderback y las notas en una asignatura de estadística de 0.14; Schutz, Drogosz, White y Distefano (1998) informan correlaciones de 0.21 y 0.16, entre las notas y SATS-Afecto y SATS-Valor, respectivamente.

Aunque en la literatura se han propuesto diferentes variables moderadoras de esta relación (véase Fitzgerald, 1997), la variable que parece tener un efecto moderador más patente es el tiempo transcurrido entre la administración del cuestionario y la realización del examen de la asignatura, de manera que cuanto mayor sea este tiempo, menor es la correlación encontrada. Este resultado se obtiene en varios trabajos de investigación, independientemente del cuestionario usado: RMARS (Harvey, Plake y Wise, 1985), SAS (Mejía, 1995; Roberts y Saxe, 1982), ATS (Green, 1994; Shultz y Koshino, 1998; Waters et al., 1988a) o SATS (Finney y Schraw, 2003).

Por otra parte, existen varios trabajos en los que las actitudes o la ansiedad hacia la estadística forman parte de un modelo multivariado para predecir el rendimiento de los alumnos, aunque sus resultados no permiten extraer conclusiones definitivas acerca del papel que juegan las actitudes o la ansiedad. La variedad de los modelos probados hace difícil la comparación entre ellos, aunque una característica común en muchos de estos estudios es que la importancia de la formación matemático-estadística en la predicción del rendimiento parece ser mayor que la ansiedad o actitud de los alumnos (Musch y Bröder, 1999; Schutz et al., 1998; Sorge y Schau, 2002). Además, aunque existen estudios en los que las actitudes o la ansiedad tienen efectos significativos importantes sobre el rendimiento (Fitzgerald et al., 1996; Nasser, 1999; Onwuegbuzie, Slate, Paterson, Watson y Schwartz, 2000), en otros trabajos esos efectos son poco importantes e incluso no significativos (Cashin, 2001; Scott, 2002; Wisenbaker, Nasser y Scott, 1999).

Evidencias de validez concurrente

En varios de los trabajos revisados se ha usado más de uno de los cuestionarios de actitudes o ansiedad hacia la estadística para intentar determinar la convergencia entre medidas. En general, se espera una correlación positiva entre cuestionarios de actitudes o entre cuestionarios de ansiedad y negativa entre cuestionarios de actitud y ansiedad. Las correlaciones encontradas en la mayoría de estos trabajos son bastante elevadas y del signo esperado, indicando una alta convergencia entre las distintas medidas.

Roberts y Reese (1987) fueron los primeros en usar conjuntamente el SAS y el ATS. En una muestra compuesta por 280 alumnos universitarios obtuvieron una correlación entre ambos cuestionarios de 0.88, lo cual interpretaron como una evidencia de que ambos cuestionarios pueden ser entendidos como formas alternativas de la misma prueba. Waters et al. (1988a) por su parte informan de una correlación de 0.73 entre el SAS y la subescala ATS-Asignatura, y de una correlación de 0.83 entre el SAS y la subescala ATS-Campo.

Schau et al. (1995) calcularon las correlaciones entre las diversas escalas del SATS y el ATS usando una amplia muestra de alumnos universitarios, en las que se puede destacar la alta relación entre la subescala ATS-Asignatura y las subescalas SATS-Afecto y SATS-Competencia Cognitiva ($r = 0.79$ y $r = 0.76$, respectivamente) y entre ATS-Campo y SATS-Valor ($r = 0.76$). Cashin y Elmore

(2000), por su parte, han utilizado conjuntamente el SAS, ATS y SATS obteniendo correlaciones significativas entre los tres instrumentos.

En el trabajo de Harvey et al. (1985) se usó un cuestionario de actitudes hacia la estadística (ATS) junto a un cuestionario de ansiedad hacia la estadística (RMARS) en dos muestras distintas. En la primera muestra obtuvieron una correlación entre ambos instrumentos de 0.06. En la segunda muestra, la correlación entre el RMARS aplicado al principio del curso y la subescala ATS-Campo fue de -0.12 y con la subescala ATS-Asignatura fue de -0.75; mientras que esas correlaciones fueron respectivamente de -0.03 y -0.60 cuando la medida del RMARS se obtuvo en un momento posterior.

Recientemente, Watson et al. (2003) aplicaron conjuntamente el SATS y el STARS a una muestra de 200 graduados universitarios matriculados en Facultades de Educación. La correlación entre las puntuaciones totales del SATS y del STARS fue de -0.89. Además se obtuvieron correlaciones entre la puntuación total en el SATS y cada una de las seis subescalas del STARS que iban desde -0.38 a -0.87.

Por último, Auzmendi (1991) para validar su Escala de Actitudes hacia la Estadística aplicó ese cuestionario junto al SAS en una muestra de 101 estudiantes de Psicología y Pedagogía, obteniendo una correlación entre ambos instrumentos de 0.86.

Evidencias de validez retrospectiva

En este apartado vamos a abordar aquellos estudios en que las actitudes hacia la estadística funcionan como variable dependiente. Dependiendo del tipo de variables predictoras hemos clasificado los resultados de estos estudios en aquellos en los que se analiza la capacidad predictiva de determinadas variables personales como el género y características de personalidad, variables relacionadas con la experiencia formativa, y variables que describen los pensamientos auto-referidos relativos a las propias capacidades.

Variables personales

A pesar de que existe cierta tradición de considerar que las mujeres tienen una mayor ansiedad o actitudes más negativas hacia la estadística que los hombres, heredada de la literatura sobre actitudes hacia las matemáticas, en la revisión realizada hemos encontrado resultados contradictorios. Algunos de los trabajos en los que se encuentran diferencias estadísticamente significativas entre hombres y mujeres son: Onwuegbuzie (1995), usando el STARS; Roberts y Saxe (1982) y Cuesta et al. (2001), usando el SAS; y Zeidner (1991) y Sigurdsson (1991), usando versiones del MARS. Mientras Onwuegbuzie (1995), Roberts y Saxe (1982) y Sigurdsson (1991) encuentran que las diferencias son las esperadas, Zeidner (1991) encuentra un patrón más complejo, según el cual las mujeres tenían una mayor ansiedad hacia el examen de estadística que los hombres y éstos tenían niveles ligeramente superiores de ansiedad hacia el contenido de la estadística. En otros trabajos el género de los alumnos no produce diferencias significativas en las puntuaciones de los cuestionarios. Así, por ejemplo, ocurre en algunos trabajos en los que se usa el ATS (Araki y Shultz, 1995; Cashin, 2001; Gil, 1999; Harvey et al., 1985), el RMARS (Harvey et al., 1985), el STARS (Baloglu, 2003; Tomazic y Katz, 1988), o el SATS (Sorge, Schau, Hubele y Kennedy, 2000).

En cuanto a la edad de los estudiantes, de nuevo se han encontrado resultados contradictorios. Con respecto al SAS de Roberts y Bilderback, se han encontrado correlaciones consistentemente bajas con la edad, aunque en algún caso esa correlación tenía un valor positivo de 0.15 (Roberts y Saxe, 1982), mientras en otro su valor era negativo de -0.14 (Roberts y Reese, 1987). Usando un cuestionario de ansiedad como el STARS también se obtienen resultados contrapuestos: mientras Tomazic y Katz (1988), usando un MANOVA con las subescalas del STARS como variables dependientes, no encuentran diferencias significativas entre estas subescalas en función de la edad; Baloglu (2003), usando un MANCOVA para determinar la influencia de la edad y el género sobre las puntuaciones en las subescalas del STARS, y controlando el efecto de la experiencia matemática previa, encuentra que los estudiantes mayores tenían niveles más altos de ansiedad hacia el examen y las clases de estadística que los estudiantes más jóvenes.

Onwuegbuzie y sus colaboradores han realizado en la última década varios estudios en los que analizan la relación entre varias características personales y la ansiedad hacia la estadística, medida a

través del STARS. En un estudio con 90 alumnos de ciencias sociales y conductuales matriculados en asignaturas de carácter metodológico, Onwuegbuzie y Daley (1997) encontraron mayores niveles de ansiedad en aquellos alumnos menos orientados hacia una inteligencia lingüística y lógico-matemática y más orientados hacia una inteligencia espacial e interpersonal. Por otra parte, Onwuegbuzie (1999) encontró que los alumnos de raza negra tienen niveles de ansiedad estadísticamente superiores a los mostrados por alumnos de raza blanca. En un estudio realizado por Onwuegbuzie y Daley (1999) para estudiar la relación entre el perfeccionismo y la ansiedad hacia la estadística, encuentran que aquellos alumnos que mantienen expectativas no realistas acerca de otras personas (perfeccionistas orientados hacia otros) y aquellos que tienen una necesidad excesiva de alcanzar objetivos prescritos por otras personas (perfeccionistas socialmente determinados) suelen tener niveles superiores de ansiedad hacia la estadística. Tenemos que señalar, sin embargo, que este último resultado ha sido puesto en cuestión recientemente por un estudio de Walsh y Ugumba-Agwunobi (2002) en el que se encuentra que la ansiedad hacia la estadística está más asociada al perfeccionismo orientado hacia uno mismo que al perfeccionismo orientado a otros.

Experiencia formativa

En la literatura especializada se ha supuesto que la experiencia formativa con asignaturas de matemáticas y de estadística es el principal determinante de las actitudes de los alumnos. En las investigaciones realizadas al respecto la experiencia previa ha sido operativizada por medio de variables como el número de asignaturas de matemáticas y de estadística cursadas, las notas obtenidas en estas asignaturas, y los conocimientos en matemáticas o estadística medidos a través de algún examen inicial.

Teóricamente es esperable que cuanto mayor sea la experiencia formativa previa en matemáticas o estadística –número de asignaturas- o mayor haya sido el aprovechamiento de la misma –notas o conocimientos-, más positiva sea la actitud y menor la ansiedad de los alumnos. Los resultados de la mayoría de las investigaciones confirman esta hipótesis, aunque los resultados son más claros cuando la relación se establece con las notas o los conocimientos que con el número de asignaturas.

Se encuentran también correlaciones significativas del número de asignaturas de matemáticas con las puntuaciones del SAS (Roberts y Saxe, 1982) y del RMARS (Royse y Rompf, 1992). Perney y Ravid (1990) obtuvieron puntuaciones significativamente más altas en ambas subescalas del ATS en aquellos alumnos que tenían una formación matemática más extensa que en otros con una menor formación. Onwuegbuzie (1998) por otra parte encontró correlaciones significativas entre el número de asignaturas de estadística cursadas con anterioridad y las puntuaciones en ATS-Campo y ATS-Asignatura. En el trabajo de Roberts y Reese (1987), sin embargo, se obtuvieron correlaciones muy bajas y no significativas entre el número de asignaturas de estadística y el número de asignaturas de matemáticas, por una parte, y las puntuaciones en el ATS y en SAS por otra.

En la mayoría de las investigaciones analizadas se han encontrado correlaciones significativas entre las puntuaciones en los cuestionarios de actitudes o ansiedad hacia la estadística y alguna medida de los conocimientos de matemáticas al empezar el curso. Se encuentran esas correlaciones significativas cuando se usan las puntuaciones en el SAS (Roberts y Saxe, 1982), el ATS (Harvey et al., 1985; Kottke, 2000; Woehlke, 1991), el RMARS (Harvey et al., 1985; Plake y Parker, 1982) y el SATS (Schutz et al., 1998). El promedio de estas correlaciones, tomando los valores absolutos de las mismas, es igual a 0.36, bastante superior al promedio de 0.19 obtenido cuando se consideran las correlaciones entre puntuaciones en los cuestionarios y número de asignaturas de matemáticas o estadística. Tomazic y Katz (1988), por otra parte, encontraron diferencias significativas en las puntuaciones de las seis subescalas del STARS entre los cuatro grupos de alumnos formados al dividir la muestra en función de que hubieran sacado buenas notas o no en matemáticas y de que hubieran cursado recientemente o no las asignaturas de matemáticas. Royse y Rompf (1992), por último, obtuvieron menos ansiedad, medida a través del RMARS, en los alumnos de su muestra que no habían suspendido nunca una asignatura de matemáticas que en aquellos que sí habían suspendido.

Cuando se ha estudiado si el conocimiento previo en estadística está relacionado con las actitudes, se han obtenido correlaciones significativas de promedio 0.30 con el SAS (Roberts y Saxe, 1982) y el ATS (Kottke, 2000). En Schutz et al. (1998), sin embargo, se encuentran correlaciones prácticamente

nulas con el SATS-Afecto y el SATS-Valor, lo cual puede ser debido a que se usó un tipo de prueba atípica para medir los conocimientos de estadística consistente en rellenar los vacíos de un árbol conceptual.

En el ámbito español, tanto Gil (1999) usando el ATS como Cuesta et al. (2001) con el SAS han obtenido relación entre las actitudes hacia la estadística y el tipo de bachillerato cursado. Concretamente, en Gil, los alumnos que habían cursado bachillerato de ciencias, con una mayor carga en asignaturas de matemáticas, tenían actitudes más positivas que los alumnos que habían cursado bachillerato de letras.

Una cuestión de investigación diferente se refiere no a la formación previa sino al efecto de determinadas características de la asignatura de estadística sobre las actitudes o la ansiedad de los alumnos. Para estudiar esta cuestión ha sido habitual aplicar el cuestionario elegido tanto al inicio como al final del curso y tomar la diferencia entre ambas medidas como un indicador del efecto de la intervención. Los resultados de estas investigaciones apuntan bien a la inexistencia de cambios en las actitudes o la ansiedad de los alumnos (Elmore, Lewis & Bay, 1993; Sgoutas-Emch y Johnson, 1998; Shultz y Koshino, 1998), la existencia de cambios sólo en algunos aspectos de la ansiedad o de las actitudes (Elmore y Lewis, 1991; Sorge et al., 2000; Waters et al., 1988a) o a la existencia de mejoras en la actitud o reducción de la ansiedad (Forte, Healey y Campbell, 1994; Harvey et al., 1985; Tomazic y Katz, 1988; Perney y Ravid, 1990; Quilter y Chester, 2001; Roberts y Saxe, 1982).

En la mayoría de estas investigaciones la existencia de cambios en las actitudes se atribuye al hecho de haber cursado la asignatura de estadística (Harvey et al., 1985; Perney y Ravid, 1990; Roberts y Saxe, 1982; Sorge et al., 2000; Waters et al., 1988a) o a alguna característica de la impartición de las mismas, como el uso de un determinado tipo de software (Forte et al., 1994), el uso de un aula de informática (Elmore y Lewis, 1991), el uso de conferencias a través de la web (Quilter y Chester, 2001) o el uso de una determinada aproximación conceptual –aproximación no cuantitativa a la enseñanza de la estadística- (Tomazic y Katz, 1988).

Aunque en todos estos trabajos, los diseños de investigación incluían la aplicación a todos los alumnos de una medida al inicio y otra al final del curso; existen tres trabajos en los que además se usó un grupo control (Elmore y Lewis, 1991; Elmore et al., 1993; Sgoutas-Emch y Johnson, 1998).

Pensamientos auto-referidos

Por último, algunas de las investigaciones revisadas sugieren que los pensamientos referidos a las propias capacidades, sobre todo las capacidades relacionadas con las matemáticas y la estadística, condicionan las actitudes y la ansiedad de los alumnos. Los resultados de los trabajos revisados muestran correlaciones moderadas e incluso altas entre ambos tipos de variables. Así, por ejemplo, Zeidner (1991) encuentra una correlación inversa significativa entre la ansiedad hacia la estadística medida con el SAI y las auto-percepciones acerca de las habilidades matemáticas ($r = -0.38$). Perney y Ravid (1990), por su parte, hallaron correlaciones desde 0.24 a 0.74 entre una escala de autoconcepto en matemáticas y las subescalas del ATS medidas al inicio y al final del curso.

La relación entre el autoconcepto en matemáticas y la ansiedad hacia la estadística ha sido estudiada también en Benson (1989) y en Bandalos et al. (1995), usando una escala elaborada por Benson para medir el autoconcepto y el TAI de Spielberger para medir la ansiedad hacia los exámenes de estadística, y modelos de ecuaciones estructurales en los que se integran ambas variables. En estas investigaciones se encuentra que cuanto más positivo sea el autoconcepto en matemáticas menor será la ansiedad hacia el examen de estadística, efecto que se observa en las puntuaciones en el TAI (Benson, 1989) y las puntuaciones en el TAI-Emotividad (Bandalos et al., 1995). En otro trabajo, Onwuegbuzie (2000a) usó un análisis de correlación canónica para estudiar qué combinación de las cinco dimensiones de la autopercepción relativas a las competencias académicas, medidas a través del SPPCS de Neemann y Harter (1986), podría estar correlacionada con alguna combinación de las dimensiones de la ansiedad hacia la estadística medida con el STARS. Los resultados de este análisis muestran que las dimensiones de la auto-percepción que permiten predecir mejor las seis dimensiones del STARS son, por este orden, la Habilidad intelectual percibida, la Competencia escolar percibida y la Creatividad percibida.

En cuanto a la autopercepción de las capacidades estadísticas, Schutz et al. (1998) preguntaron a sus alumnos acerca de la confianza que tenían en dominar el material de introducción a la estadística y encontraron correlaciones significativas entre las respuestas a esta pregunta y las subescalas SATS-Afecto ($r = 0.51$) y SATS-Valor ($r = 0.44$). Silva, Cazorla y Brito (1999), por su parte, solicitaron a los alumnos que clasificaran su conocimiento de estadística, optando entre ‘pésimo’, ‘pasable’ y ‘bueno’. Posteriormente usaron un análisis de varianza con los resultados de esa pregunta como medida de la autopercepción acerca de la capacidad estadística y las puntuaciones en la *Escala de actitudes em relação à Estatística* de Cazorla como variable dependiente. El resultado de este análisis mostró diferencias significativas entre los tres grupos, con actitudes más positivas para aquellos que consideraban tener más conocimientos de estadística.

Finney y Schraw (2003) han desarrollado medidas específicas de la autoeficacia de los alumnos en estadística y las han aplicado a una muestra de alumnos junto a otros cuestionarios, incluido el SATS de Schau et al. Los resultados de este estudio muestran correlaciones de las subescalas del SATS con una medida de la confianza en las propias capacidades para aprender las destrezas necesarias para resolver tareas específicas relacionadas con la estadística (SELS) que van desde 0.37 a 0.65.

6. CONCLUSIONES DE LA REVISIÓN REALIZADA

A pesar de su importancia para el proceso de enseñanza de la estadística, las investigaciones sobre las actitudes y la ansiedad hacia la estadística han sido y siguen siendo más escasas que las que se dan en campos afines, como el estudio de las actitudes hacia la ciencia o de las actitudes hacia las matemáticas (Gal et al., 1997). Esto no ha sido óbice, sin embargo, para que haya existido una cierta proliferación de instrumentos de medida que, en la mayoría de los casos, son usados en muy pocas investigaciones. Una de las consecuencias de este hecho es que no existen suficientes evidencias de la validez de muchos de los instrumentos de medida de las actitudes o la ansiedad hacia la estadística existentes.

En la revisión realizada hemos encontrado 17 instrumentos para medir actitudes o ansiedad hacia la estadística. Todos ellos son cuestionarios compuestos por ítems con opciones de respuesta propias de las escalas de valoración, habitualmente formatos de respuesta tipo Likert. Los cuestionarios de actitudes más usados son, por este orden, el ATS de Wise, el SATS de Schau y el SAS de Roberts. En cuanto a los cuestionarios de ansiedad hacia la estadística los más usados son el STARS de Cruise, seguido de la versión del MARS de Suinn elaborada por Plake y Parker. Los resultados más destacados de las investigaciones en las que se estudian las propiedades psicométricas de estos cuestionarios son las siguientes:

- a) La mayoría de las escalas que componen los cuestionarios analizados han mostrado una alta fiabilidad de consistencia interna. Por otra parte, sólo existen dos estudios de la fiabilidad entendida como estabilidad temporal de las puntuaciones, obteniéndose en ambos casos valores aceptables del coeficiente de estabilidad, para el ATS en un caso y el STARS en el otro.
- b) En cuanto a las evidencias de validez basadas en la estructura interna, la estructura del dominio de las actitudes hacia la estadística ha sido concebida teóricamente de muy diversas formas, desde una estructura unidimensional en el SAS de Roberts hasta una compuesta por cinco dimensiones en la EAE de Auzmendi. Lo mismo se puede decir del dominio de la ansiedad hacia la estadística, con propuestas que van desde la única dimensión del SAS de Pretorius y Norman hasta las seis dimensiones del STARS de Cruise et al. En todos los casos en los que se ha investigado si las respuestas de los sujetos a los cuestionarios respectivos se ajustaban a estas diferentes estructuras se han usado técnicas de análisis factorial. Los resultados de muchas de estas investigaciones apoyan las estructuras preconcebidas, siendo destacables las evidencias a favor de la estructura bidimensional del ATS obtenidas con técnicas de análisis factorial exploratorio y las evidencias a favor de la estructura de cuatro dimensiones del SATS obtenidas con técnicas de análisis factorial confirmatorio.

- c) En cuanto a las evidencias de validez del contenido, es destacable que para la mayoría de los cuestionarios analizados no se aporten este tipo de evidencias. Sólo con cuatro cuestionarios (STARS, ATS, SATS y QAQ) se realizan estudios para investigar el contenido de sus respectivos ítems a través de las evaluaciones de jueces expertos. Estas evaluaciones, básicamente juicios acerca de la congruencia ítem-dimensión, se realizan en las fases iniciales de desarrollo de los cuestionarios y sirven fundamentalmente como un método de selección de ítems.
- d) En cuanto a las evidencias de validez basadas en las relaciones con otras variables, cuando las puntuaciones en los cuestionarios se han usado para predecir el rendimiento de los alumnos en las asignaturas de estadística se ha encontrado que ambas variables están relacionadas débilmente. En cambio, en aquellos estudios en que se han usado más de uno de los cuestionarios revisados se ha hallado por lo general una alta convergencia entre ellos. Con respecto a las variables que permiten predecir las actitudes o la ansiedad de los alumnos, tanto la formación previa en matemáticas y estadística como el autoconcepto acerca de las capacidades relacionadas con estas materias han mostrado relaciones consistentes con las puntuaciones en los cuestionarios de actitudes o ansiedad. Los resultados son menos claros cuando se usa el género de los alumnos, su edad u otras características personales para intentar predecir las actitudes.

A partir de esta revisión de las evidencias de validez es posible concluir que algunos de estos cuestionarios, sobre todo el ATS y el SATS, pueden ser usados con ciertas garantías para evaluar las actitudes hacia la estadística. De igual forma, consideramos que, en líneas generales, el STARS es un cuestionario adecuado para medir ansiedad hacia la estadística. Sin embargo, también se han detectado algunas carencias en el proceso de elaboración y validación de estos instrumentos que hacen recomendables nuevas investigaciones para el desarrollo de instrumentos de medida de las actitudes y la ansiedad hacia la estadística o para completar el análisis de los existentes con suficientes garantías de validez.

En este sentido, estamos de acuerdo con Gal et al. (1997) cuando afirmaban que uno de los principales problemas de la investigación en este campo es la falta de modelos teóricos que guíen los trabajos. Esta falta de fundamentación teórica se refleja, por ejemplo, en uno de los aspectos más importantes para la elaboración y validación de un instrumento de medida, a saber, la determinación de su estructura interna. En la práctica, los autores de cada uno de los cuestionarios revisados han propuesto estructuras diferentes del dominio de las actitudes hacia la estadística. Sin embargo, casi ninguno de estos autores ha justificado la elección de esas estructuras. En aquellos casos en los que se ha propuesto alguna justificación, por ejemplo en el SATS, éstas suelen estar basadas en el conocimiento, no necesariamente bien informado, de profesores y alumnos de estadística acerca de ese dominio conductual más que en una teoría acerca del mismo. Otra consecuencia de la falta de teorías se refleja en la dificultad para determinar claramente una red nomológica de relaciones entre las actitudes hacia la estadística y otras variables. Así, por ejemplo, en cuanto a la relación entre género y actitudes de los alumnos, no está claro si es esperable que los alumnos tengan actitudes más, menos o igual de favorables hacia la estadística que las alumnas.

Además de la falta de modelos teóricos, existen otras carencias más específicas de los estudios de elaboración y validación de estos instrumentos, algunas de las cuales comentamos a continuación. En primer lugar, puesto que muchos de estos cuestionarios se van a usar en distintos momentos del curso para estudiar la evolución de las actitudes o la ansiedad de los alumnos, es bastante conveniente estudiar la estabilidad de las puntuaciones, cosa que sólo se ha hecho en el caso del ATS y el STARS. Por otra parte, faltan estudios que aporten evidencias basadas en el contenido, y cuando éstos se realizan se usan tareas clásicas de evaluación del contenido que han sido cuestionadas debido a que a la información previa que se proporciona a los jueces condiciona en gran medida sus juicios (Sireci, 1998). Un último aspecto de carácter técnico que, desde nuestro punto de vista, constituye una limitación de estos estudios de validación es el uso casi exclusivo de técnicas de análisis factorial exploratorio en el estudio de la estructura interna de los cuestionarios, con la excepción del SATS.

A pesar de estas limitaciones, el análisis de los trabajos revisados permite extraer algunas recomendaciones para futuras investigaciones en este campo:

En primer lugar, consideramos que la estructura de cuatro dimensiones propuesta para el SATS es una candidata plausible para describir el dominio de las actitudes hacia la estadística. Por otra parte, no existen evidencias suficientes que permitan determinar si una estructura de dos dimensiones (Plake y Parker, 1982; Zeidner, 1991) para el dominio de la ansiedad hacia la estadística es más adecuada que una estructura de seis dimensiones (Cruise et al., 1985). Por ello deberían diseñarse investigaciones para poner a prueba diferentes estructuras internas para este dominio desde una óptica confirmatoria, en la línea de lo realizado por los autores del SATS (Schau et al., 1995; Dauphinee et al., 1997). Además, en los últimos años se ha propuesto una interesante alternativa para el estudio de la estructura de dominios conductuales (Deville, 1996), que creemos tiene gran potencial en el estudio de las actitudes y la ansiedad hacia la estadística. Se trata de una propuesta que pone el foco en una especificación clara del contenido, una redacción de ítems que se ajuste a esas especificaciones y el uso de técnicas de análisis de datos que permitan el análisis conjunto de las evidencias relativas a la estructura interna y al contenido. Tal como recomiendan los *Standards for Educational and Psychological Testing* (AERA/APA/NCME, 1999) en esta propuesta se integran diferentes tipos de evidencias de validez, lo cual viene además a subsanar una de las carencias detectadas en la revisión realizada, a saber, la escasez de estudios de validación de contenido.

En cuanto a la relación de las actitudes y la ansiedad hacia la estadística con el rendimiento de los alumnos en las asignaturas de estadística, estos constructos suelen mostrar una capacidad muy limitada para predecir el rendimiento. En los últimos años se han propuesto, sin embargo, interesantes modelos multivariados para la predicción del rendimiento en asignaturas de estadística (Lalonde y Gardner, 1993; Onwuegbuzie, 2003; Tremblay, Gardner y Heipel, 2000) que deberían ser puestos a prueba en nuevos estudios. Otra prometedora línea de trabajo en este ámbito, que sin embargo no ha sido investigada hasta el momento, es el estudio de la relación de las actitudes y la ansiedad hacia la estadística con algunas variables relacionadas con el rendimiento en estadística, como el abandono de las asignaturas optativas o la falta de asistencia a las convocatorias de examen. Los resultados del trabajo de Walsh y Ugumba-Agwunobi (2002) que muestran una relación significativa entre la ansiedad hacia la estadística y una cierta tendencia a retrasar la realización de tareas, parecen avalar la conveniencia de realizar dichos estudios acerca de algunas posibles manifestaciones conductuales de esta tendencia.

Por último, es necesario destacar que en la mayoría de trabajos analizados se ha encontrado que la formación previa en estadística y matemáticas parece ser un antecedente bastante consistente de las actitudes y la ansiedad hacia la estadística. Esta relación, sin embargo, tiende a ser más débil que la encontrada entre las actitudes o la ansiedad con indicadores de la percepción que tienen los alumnos acerca de sus capacidades en estas materias (Schutz et al., 1998). Por ello consideramos conveniente que en los modelos de trabajo de futuras investigaciones se incorporen medidas específicas de este aspecto del autoconcepto, como las medidas de autoeficacia en estadística desarrolladas por Finney y Schraw (2003).

REFERENCIAS

- Aiken, L. R., y Dreger, R. M. (1961). The effect of attitudes in performance in mathematics. *Journal of Educational Psychology*, 52(1), 19-24.
- American Educational Research Association, the American Psychological Association y the National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington, DC, EE. UU.: AERA.
- Araki, L. T., y Shultz, K. S. (1995, Abril). *Student attitudes toward statistics and their retention of statistical concepts*. Comunicación presentada al Annual Meeting of the Western Psychological Association, Los Angeles.
- Auzmendi, E. (1991). *Evaluación de las actitudes hacia la estadística en estudiantes universitarios y factores que las determinan*. Tesis doctoral no publicada, Universidad de Deusto, Bilbao, España.

- Auzmendi, E. (1992). *Las actitudes hacia la matemática-estadística en las enseñanzas medias y universitaria*. Bilbao, España: Ediciones Mensajero.
- Baloglu, M. (2002). Psychometric properties of the Statistics Anxiety Rating Scale. *Psychological Reports, 90*, 315-325.
- Baloglu, M. (2003). Individual differences in statistics anxiety among collage students. *Personality and Individual Differences, 34*, 855-865.
- Bandalos, D. L., Yates, K., y Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on test anxiety. *Journal of Educational Psychology, 87*(4), 611-623.
- Bending, A. W., y Hughes, J. B. (1954). Student attitude and achievement in a course in introductory statistics. *Journal of Educational Psychology, 45*, 268-276.
- Benson, J. (1989). Structural components of statistical test anxiety in adults: An exploratory model. *Journal of Experimental Education, 57*, 247-261.
- Bessant, K. C. (1997). The development and validation of scores on the Mathematics Information Processing Scale (MIPS). *Educational and Psychological Measurement, 57*(5), 841-857.
- Betz, N. E. (1978). Prevalence, distribution and correlates of math anxiety in college students. *Journal of Counseling Psychology, 25*, 441-448.
- Birenbaum, M., y Eylath, S. (1994). Who is afraid of statistics? Correlates of statistics anxiety among students of educational sciences. *Educational Research, 36*(1), 93-98.
- Carmona, J., y Moreno, R. (1999). Structuring the domain of attitudes toward statistics. En R. Meyer Schweizer (Ed.), *Facet Theory Conference: Design and Analysis* (pp. 323-334). Berna, Suiza: FTA-c/o Institut für Soziologie.
- Cashin, S. E. (2001). *Effects of mathematics self-concept, perceived self-efficacy, and attitudes toward statistics on statistics achievement*. Tesis doctoral no publicada, Southern Illinois University, Carbondale, IL, EE. UU.
- Cashin, S. E., y Elmore, P. B. (1997, Marzo). *Instruments used to assess attitudes toward statistics: A psychometric evaluation*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Chicago.
- Cashin, S. E., y Elmore, P. B. (2000, Octubre). *The Survey of Attitudes Toward Statistics scale: A construct validity study*. Comunicación presentada al Annual Meeting of the Mid-Western Educational Research Association, Chicago, IL, EE. UU.
- Cazorla, I. M., Silva, C. B., Vendramini, C., y Brito, M. R. F. (1999). *Adaptação e validação de uma Escala de Atitudes em Relação à Estatística*. Comunicación presentada a la Conferência Internacional Experiências e Expectativas do Ensino da Estatística: Desafios para o Século XXI, Florianópolis, Brasil. [Online: <http://www.inf.ufsc.br/cee/pasta1/art2.html>].
- Chang, L. (1996). Quantitative Attitudes Questionnaire: Instrument development and validation. *Educational and Psychological Measurement, 56*(6), 1037-1042.
- Cherian, V. I., y Glencross, M. J. (1997). Sex, socio-economic status, and attitude toward applied statistics among postgraduate education students. *Psychological Reports, 80*, 1385-1386.
- Crocker, L., y Algina, J. (1986). *Introduction to classical and modern test theory*. Nueva York: Holt, Rinehart & Winston.
- Cruise, R. J., Cash, R. W., y Bolton, D. L. (1985). Development and validation of an instrument to measure statistical anxiety. *ASA Proceedings of the Section on Statistical Education* (pp. 92-97). Alexandria, VA, EE. UU.: American Statistical Association.
- Cuesta, M., Rifá, H., y Herrero, F. J. (2001, Septiembre). *Un estudio exploratorio, en estudiantes de psicología, de una escala de actitudes hacia la estadística*. Póster presentado en el VII Congreso de Metodología de las Ciencias Sociales y de la Salud, Madrid. Resumen descargado el 6 de Marzo de 2002 de http://www.uned.es/VIIICongresoMetodologia/poster/sesionI_titulos.htm

- Dauphinee, T. L., Schau, C., y Stevens, J. J. (1997). Survey of Attitudes Toward Statistics: Factor structure and factorial invariance for women and men. *Structural Equation Modelling*, 4(2), 129-141.
- Deville, C. W. (1996). An empirical link of content and construct validity evidence. *Applied Psychological Measurement*, 20(2), 127-139.
- Dillon, K. M. (1982). Statisticophobia. *Teaching of Psychology*, 9(2), 117.
- Dutton, W. A. (1951). Attitudes of prospective teachers toward mathematics. *Elementary School Journal*, 52, 84-90.
- Eagly, A. H., y Chaiken, S. (1998). Attitude structure and function. En D. T. Gilbert, S. T. Fiske y G. Lindzey (Eds.), *The handbook of social psychology* (4ª Ed, Vol. 1, pp. 269-322). Nueva York, NY, EE. UU.: McGraw-Hill.
- Elmore, P. B., y Lewis, E. L. (1991, Abril). *Statistics and computer attitudes and achievement of students enrolled in Applied Statistics: Effect of a computer laboratory*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Chicago.
- Elmore, P. B., Lewis, E. L., y Bay, M. L. G. (1993). *Statistics achievement: A function of attitudes and related experiences*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Atlanta, GA, EE. UU. (ERIC Document Reproduction Service No. ED360324)
- Estrada, A. (2001). *Actitudes hacia la estadística e instrumentos de evaluación*. Comunicación presentada a las Jornadas Europeas de Estadística: La Enseñanza y la Difusión de la Estadística. Palma de Mallorca, España.
[Online: http://www.caib.es/ibae/esdeveniment/jornades_10_01/doc/estadistica.doc]
- Estrada, A. (2002). *Análisis de las actitudes y conocimientos estadísticos elementales en la formación del profesorado*. Tesis doctoral no publicada, Universidad Autónoma de Barcelona, España.
- Finney, S. J., y Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology*, 28, 161-186.
- Fitzgerald, S. M. (1997). *The relationship between anxiety and statistics achievement: A meta-analysis*. Tesis doctoral no publicada, University of Toledo, Toledo, OH, EE. UU.
- Fitzgerald, S. M., Jurs, S. J., y Hudson, L. M. (1996). A model predicting statistics achievement among graduate students. *College Student Journal*, 30(3), 361-366.
- Forte, J. A., Healey, J., y Campbell, M. H. (1994). Does MicroCase statistical software package increase the competence and comfort of undergraduate social work and social science majors? *Journal of Teaching in Social Work*, 10(1/2), 99-115.
- Gal, I. y Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Toward an assessment framework. *Journal of Statistics Education*, 2(2). Descargado el 6 de Marzo de 2003 de: <http://www.amstat.org/publications/jse/v2n2/gal.html>
- Gal, I., Ginsburg, L., y Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. En I. Gal y J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 37-51). Netherlands: IOS Press.
- Gil, J. (1999). Actitudes hacia la estadística. Incidencia de las variables sexo y formación previa. *Revista Española de Pedagogía*, 214, 567-590.
- Glencross, M. J., y Cherian, V. J. (1992). Attitudes toward applied statistics of postgraduate education students in Transkei. *Psychological Reports*, 70, 67-75.
- Green, K. E. (1993, Abril). *Affective, evaluative and behavioral components of attitudes toward statistics*. Comunicación presentada en Annual Meeting of the American Educational Research Association, Atlanta, GA, EE. UU.
- Green, K. E. (1994, Abril). *The affective component of attitude in statistics instruction*. Comunicación presentada en Annual Meeting of the American Educational Research Association, Nueva Orleans, LA, EE. UU.

- Harvey, A. L., Plake, B. S., y Wise, S. L. (1985). *The validity of six beliefs about factors related to statistics achievement*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Chicago, IL, EE. UU. (ERIC Document Reproduction Service No. ED262965)
- Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development*, 34(3), 177-189.
- Huedo, T., López, J. A., Martínez, R., y Nortes, A. (2003). *Contenidos y actitudes en estadística: Un estudio en maestros en formación*. Comunicación presentada al 27 Congreso Nacional de Estadística e Investigación Operativa, Lleida, España.
[Online: http://www.udl.es/usuarios/seio2003/treballs/05_1_5.pdf]
- Kottke, J. L. (2000). Mathematical proficiency, statistics knowledge, attitudes toward statistics, and measurement course performance. *College Student Journal*, 34(3), 334-347.
- Lalonde, R. N., y Gardner, R. C. (1993). Statistics as a second language: A model for predicting performance in psychology students. *Canadian Journal of Behavioural Science*, 25(1), 108-125.
- McCall, C. H., Belli, G., y Madjidi, F. (1991). The complexities of teaching graduate students in educational administration introductory statistical concepts. *PICTeachSt3*, 2, 495-497.
- Mejía, J. C. (1995). *Reliability and validity study of a Spanish translation of the Statistical Attitude Survey*. Tesis doctoral no publicada, Pennsylvania State University, PA, EE. UU.
- Miller, R. B., Behrens, J. T., Green, B. A., y Newman, D. (1993). Goals and perceived ability: Impact on student valuing, self-regulation and persistence. *Contemporary Educational Psychology*, 18, 2-14.
- Musch, J., y Bröder, A. (1999). Test anxiety versus academic skills: A comparison of two alternative models for predicting performance in a statistics exam. *British Journal of Educational Psychology*, 69, 105-116.
- Nasser, F. (1999). *Prediction of college students achievement in Introductory Statistics Course*. Comunicación presentada a la 52nd ISI –International Statistical Institute- Session, Helsinki, Finlandia. [Online: <http://www.stat.fi/isi99/proceedings/arkisto/varasto/nass0341.pdf>]
- Neemann, J., y Harter, S. (1986). *Manual for the self-perception profile for college students*. Manuscrito no publicado, University of Denver, CO, EE. UU.
- Onwuegbuzie, A. J. (1993). The interaction of statistics test anxiety and examination condition in statistics achievement of post baccalaureate non statistics majors. *Dissertation Abstracts International*, 54(12), 4371A.
- Onwuegbuzie, A. J. (1995). Statistics test anxiety and female students. *Psychology of Women Quarterly*, 19, 413-418.
- Onwuegbuzie, A. J. (1998). Teachers' attitudes toward statistics. *Psychological Reports*, 83, 1008-1010.
- Onwuegbuzie, A. J. (1999). Statistics anxiety among African American graduate students: An effective filter? *Journal of Black Psychology*, 25(2), 189-209.
- Onwuegbuzie, A. J. (2000a). Statistics anxiety and the role of self-perceptions. *Journal of Educational Research*, 93(5), 323-335.
- Onwuegbuzie, A. J. (2000b). Attitudes toward statistics assessments. *Assessment and Evaluation in Higher Education*, 25(4), 321-339.
- Onwuegbuzie, A. J. (2003). Modeling statistics achievement among graduate students. *Educational and Psychological Measurement*, 63(6), 1020-1038.
- Onwuegbuzie, A. J. (en prensa). Prevalence of statistics anxiety among graduate students. *Journal of Research in Education*.
- Onwuegbuzie, A. J., y Daley, C. E. (1997). The role of multiple intelligences in statistics anxiety. *Research and Reflections on Leadership and Organizational Change*, 3(2). Descargado el 28 de Octubre de 2000 de: <http://www.gonzaga.edu/rr/v3n2/onwuegbuzie.htm>

- Onwuegbuzie, A. J., y Daley, C. E. (1999). Perfectionism and statistics anxiety. *Personality and Individual Differences*, 26, 1089-1102.
- Onwuegbuzie, A. J., Slate, J., Paterson, F., Watson, M., y Schwartz, R. (2000). Factors associated with underachievement in educational research courses. *Research in the Schools*, 7, 53-65.
- Onwuegbuzie, A. J., y Wilson, V. A. (2003). Statistics anxiety: nature, etiology, antecedents, effects and treatments-a comprehensive review of the literature. *Teaching in Higher Education*, 8(2), 195-209.
- Perney, J., y Ravid, R. (1990, Abril). *The relationship between attitudes towards statistics, math self-concept, test anxiety and graduate students' achievement in an introductory statistics course*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Boston, MA, EE. UU.
- Plake, B. S., y Parker, C. S. (1982). The development and validation of a revised version of the Mathematics Anxiety Rating Scale. *Educational and Psychological Measurement*, 42, 551-557.
- Pretorius, T. B., y Norman, A. M. (1992). Psychometric data on the Statistics Anxiety Scale for a sample of south African students. *Educational and Psychological Measurement*, 52(4), 933-937.
- Quilter, S. M., y Chester, C. (2001). The relationship between web-based conferencing and instructional outcomes. *International Journal of Instructional Media*, 28(1), 13-22.
- Rhoads, T. R., y Hubele, N. F. (2000). Student attitudes toward statistics before and after a computer-integrated introductory statistics course. *IEEE Transactions on Education*, 43(2), 182-187.
- Roberts, D. M., y Bilderback, E. W. (1980). Reliability and validity of a statistics attitude survey. *Educational and Psychological Measurement*, 40, 235-238.
- Roberts, D. M., y Reese, C. M. (1987). A comparison of two scales measuring attitudes toward statistics. *Educational and Psychological Measurement*, 47, 759-764.
- Roberts, D. M., y Saxe, J. E. (1982). Validity of a Statistics Attitude Survey: A follow-up study. *Educational and Psychological Measurement*, 42, 907-912.
- Rosenthal, B. (1992). No more sadistics, no more sadists, no more victims. *UMAP Journal*, 13, 281-290.
- Royse, D., y Rompf, E. L. (1992). Math anxiety: A comparison of social work and non-social work students. *Journal of Social Work Education*, 28(3), 270-277.
- Sánchez-López, C. R. (1996). Validación y análisis ipsativo de la Escala de Actitudes hacia la Estadística (EAE). *Análisis y Modificación de Conducta*, 22, 799-819.
- Sarason, I. G. (1972). Experimental approaches to test anxiety: Attention and the assessment of information. En C. D. Spielberger (Ed.), *Anxiety: Current trends in theory and research* (vol. 2, pp. 283-403). Nueva York, NY, EE. UU.: Academic Press.
- Schau, C., Dauphinee, T. L., y Del Vecchio, A. (1993, Abril). *Evaluation of two surveys measuring students' attitudes toward statistics*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Atlanta, GA, EE. UU.
- Schau, C., Stevens, J., Dauphinee, T. L., y Del Vecchio, A. (1995). The development and validation of the Survey of Attitudes Toward Statistics. *Educational and Psychological Measurement*, 55, 868-875.
- Schuessler, K. F. (1993). Scaling social life feelings: Persistent issues. En D. Krebs y P. Schmidt (Eds.), *New directions in attitude measurement* (pp. 21-40). Berlin, Alemania: Walter de Gruyter.
- Schutz, P. A., Drogosz, L. M., White, V. E., y Distefano, C. (1998). Prior knowledge, attitude, and strategy use in an introduction to statistics course. *Learning and Individual Differences*, 10(4), 291-308.
- Scott, J. S. (2002). *Modeling aspects of students' attitudes and performance in an undergraduate introductory statistics course*. Tesis doctoral no publicada, University of Georgia, Athens, GA, EE. UU.
- Sgoutas-Emch, S. A., y Johnson, C. J. (1998). Is journal writing an effective method of reducing anxiety toward statistics?. *Journal of Instructional Psychology*, 25(1), 49-57.

- Shultz, K. S., y Koshino, H. (1998). Evidence of reliability and validity for Wise's Attitude Toward Statistics scale. *Psychological Reports*, 82, 27-31.
- Sigurdsson, J. F. (1991). Computer experience, attitudes toward computers and personality characteristics in psychology undergraduates. *Personality and Individual Differences*, 12(6), 617-624.
- Silva, C. B., Cazorla, I. M., y Brito, M. R. F. (1999). *Concepções e atitudes em relação à estatística*. Comunicación presentada a la Conferência Internacional Experiências e Expectativas do Ensino da Estatística: Desafios para o Século XXI, Florianópolis, Brasil.
[Online: <http://www.inf.ufsc.br/cee/pasta1/art2.html>]
- Sireci, S. G. (1998). Gathering and analysing content validity data. *Educational Assessment*, 5(4), 299-321.
- Sorge, C., y Schau, C. (2002). *Impact of engineering students' attitudes on achievement in statistics: A structural model*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Nueva Orleans, LA, EE. UU.
[Online: <http://www.unm.edu/~cschau/AERA2002.pdf>]
- Sorge, C., Schau, C., Hubele, N., y Kennedy, D. (2000, Agosto). *Measuring engineering students' attitudes toward statistics and probability using Internet data collection techniques*. Comunicación presentada al Joint Statistical Meetings, Indianapolis, IN, EE. UU.
- Spielberger, C. (1980). *Preliminary Professional Manual for the Test Anxiety Inventory*. Palo Alto, CA, EE. UU.: Consulting Psychologists Press.
- Suanpang, P., Petocz, P., y Kalceff, W. (2003). *Student attitudes to learning business statistics online vs traditional methods*. Comunicación presentada a HERDSA Conference 2003, Christchurch, Nueva Zelanda.
[Online: <http://surveys.canterbury.ac.nz/herdsa03/pdfsnon/N1177.pdf>]
- Suinn, R. M., Eddie, C. A., Nicoletti, J., y Spinelli, P. R. (1972). The MARS, a measure of mathematics anxiety: Psychometric data. *Journal of Clinical Psychology*, 28, 373-375.
- Sutarso, T. (1992). *Some variables in relation to students' anxiety in learning statistics*. Comunicación presentada a Annual Meeting of the Mid-South Educational Research Association, Knoxville, TN, EE. UU. (ERIC Document Reproduction Service No. ED353334).
- Tomazic, T. J., y Katz, B. M. (1988). Statistical anxiety in Introductory Applied Statistics. *ASA Proceedings of the Section on Statistical Education* (pp. 239-244). Alexandria, VA, EE. UU.: American Statistical Association.
- Tremblay, P. F., Gardner, R. C., y Heipel, G. (2000). A model of the relationship among measures of affect, aptitude, and performance in introductory statistics. *Canadian Journal of Behavioural Science*, 32(1), 40-48.
- Velandrino, A. P., y Parodi, L. H. (1999). *La Escala de Actitudes Hacia la Estadística (EAE): Desarrollo y Propiedades Psicométricas*. Comunicación presentada a la Conferência Internacional Experiências e Expectativas do Ensino da Estatística: Desafios para o Século XXI, Florianópolis, Brasil. [Online: <http://www.inf.ufsc.br/cee/comunica/EAE.html>]
- Walsh, J. J., y Ugumba-Agwunobi, G. (2002). Individual differences in statistics anxiety: the roles of perfectionism, procrastination and trait anxiety. *Personality and Individual Differences*, 33, 239-251.
- Waters. L. K., Martelli, T. A., Zakrajsek, T., y Popovich, P. M. (1988a). Attitudes toward statistics: An evaluation of multiple measures. *Educational and Psychological Measurement*, 48, 513-516.
- Waters. L. K., Martelli, T. A., Zakrajsek, T., y Popovich, P. M. (1988b). Factor analyses of two measures of attitudes toward statistics. *Educational and Psychological Measurement*, 48, 1037-1041.
- Watson, F., Kromrey, J., Ferron, J., Lang, T., y Hogarty, K. (2003). *An assessment blueprint for Encstat: A statistics anxiety intervention program*. Comunicación presentada al AERA Annual Meeting, San Diego, CA, EE. UU.

- Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement*, 45, 401-405.
- Wisembaker, J., Nasser, F., y Scott, J. S. (1999). *A cross-cultural comparison of path models relating attitudes about and achievement in Introductory Statistics Courses*. Comunicación presentada a la 52nd ISI -Internacional Statistical Institute- Session, Helsinki, Finlandia. [Online: <http://www.stat.fi/isi99/proceedings/arkisto/varasto/wise0263.pdf>]
- Woehlke, P. L. (1991). *An examination of the factor structure of Wise's Attitude Toward Statistics scale*. Comunicación presentada al Annual Meeting of the American Educational Research Association, Chicago, IL, EE. UU. (ERIC Document Reproduction Service No. ED337500)
- Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: Some interesting parallels. *British Journal of Educational Psychology*, 61(3), 319-328.

JOSÉ CARMONA MÁRQUEZ
Departamento de Psicología
Universidad de Huelva
Avda. Fuerzas Armadas s/n.
21007 Huelva, España