

INSTRUMENTOS Y MÉTODOS PARA LA EVALUACIÓN DE LAS ACTITUDES RELACIONADAS CON LA CIENCIA, LA TECNOLOGÍA Y LA SOCIEDAD

MANASSERO MAS, MARÍA ANTONIA¹ y VÁZQUEZ ALONSO, ÁNGEL²

¹ Departamento de Psicología, Universidad de las Islas Baleares. Carretera de Valldemossa, km. 7,5
07071 Palma de Mallorca
dpsamm0@ps.uib.es

² Departamento de Física, Edificio Mateu. Universidad de las Islas Baleares.
Carretera de Valldemossa, km. 7,5. 07071 Palma de Mallorca
dfsava0@ps.uib.es

Resumen. Se presenta una aplicación empírica de instrumentos y métodos para la evaluación de las actitudes relacionadas con la ciencia, la tecnología y la sociedad. El instrumento es una adaptación del cuestionario de Opiniones sobre Ciencia, Tecnología y Sociedad (VOSTS) y la metodología de valoración supone una nueva forma de aplicar y valorar las cuestiones, que tiene las ventajas de ser invariante, emplear toda la información de las cuestiones y permitir todo tipo de aplicaciones de inferencia estadística, para la investigación, y diagnóstico, para la educación de las actitudes en el aula.

Summary. A new method and its empirical application for the evaluation of the attitudes related to science, technology and society are applied to a pool of items. The assessment methodology creates a new form of answering the items, a new metric to score attitudes and a categorization of the items to apply the metric. The advantages for science education research and educational diagnostic of the attitudes within the science classroom are the invariance of the method, the use of the whole information in the items, and the application for statistical inferences.

Como consecuencia de la tradición liberal, la educación intencional (planificada) de las actitudes en la escuela encuentra serias resistencias, porque se percibe como un tipo de indoctrinación que afecta a la esfera privada, en la que la escuela, como institución, se debería mantener neutral, es decir, no debería educar en actitudes y valores. Sin embargo, aunque no exista un proyecto de educación actitudinal, es un hecho cierto que las actitudes y valores, como expresión del contexto sociocultural-familiar en el que se vive, son transmitidas incontrolada e implícitamente por la sociedad, la familia y el profesorado en la clase, y por el ambiente y las relaciones vividas en el centro escolar, a través del llamado *currículo oculto*. El papel neutral de la escuela es un mito y la educación en valores, sin caer en la indoctrinación, es una necesidad social en la realidad actual (Camps, 1994).

El currículo actual de la reforma educativa española enfatiza las actitudes (valores y normas) como contenidos del currículo escolar, diferenciados de los contenidos de conocimientos y procedimiento, sentando las bases para su educación y evaluación en la escuela. En la educación, muchas actitudes y valores tienen preferentemente un sentido moral y ético general (solidaridad, tolerancia, libertad, pacifismo, igualdad, etc.). Además, en las diferentes áreas de conocimiento también se definen contenidos actitudinales propios de cada área, aunque éstos son comúnmente percibidos por el profesorado más bien como instrumentos o causas que favorecen el aprendizaje (actitudes positivas) o que dificultan o impiden los aprendizajes (actitudes negativas) y, frecuentemente, se identifican con (des)motivación, (des)interés y esfuerzo en los estudiantes. Pero también las actitudes propias de las áreas pueden ser planteadas

como una consecuencia o efecto de la educación, es decir, como objetivos y contenidos específicos de la educación del área, que requieren planificación, aplicación y evaluación.

EVALUACIÓN DE LAS ACTITUDES RELACIONADAS CON LA CIENCIA

En el contexto educativo, el sentido más común del término *evaluación* se refiere a la evaluación por el profesorado del progreso en los aprendizajes del alumnado y suele llevar asignada una función calificadora con efectos académicos para la promoción de curso, la titulación o el acceso a otras enseñanzas. Más allá de detalles de estilo, metodología o preparación, la evaluación educativa tradicional ha consistido en comparar, para su valoración, las producciones del alumnado (escritas, orales, manuales, etc.) frente a un patrón de referencia establecido por el cuerpo de conocimientos disciplinar, que actúa como criterio último y universal. Desde esta perspectiva, la problemática de la evaluación de los aprendizajes se centra, principalmente, en la fiabilidad y validez de los instrumentos de evaluación aplicados respecto a las capacidades cognitivas que evalúan, sobre la cual existe una amplia literatura en didáctica general (Álvarez, 1985; Stufflebeam y Shinkfield, 1987; Tenbrink, 1983; Walberg y Haertel, 1990), aunque menos en el campo específico de la didáctica de las ciencias (Alonso, Gil, y Martínez, 1992; Rodríguez, Gutiérrez y Molledo, 1992; Kempa, 1985; Satterly y Swann, 1988).

La evaluación de los aprendizajes se ha complicado más con la ampliación de los contenidos curriculares procedimentales y actitudinales. Los procedimientos se perciben por el profesorado de ciencias como algo relativamente familiar en cuanto se relacionan con el trabajo experimental, práctico, de campo o de laboratorio, que siempre han sido considerados como una parte de la enseñanza de las ciencias. Sin embargo, en el caso de los contenidos actitudinales, la novedad es mayor y su evaluación, un problema adicional, pues al ser dependientes del contexto y carecer de un cuerpo de conocimientos de referencia, como en el caso de los conocimientos o procedimientos, su evaluación resulta más difícil. Los especialistas sostienen que la evaluación de las actitudes y valores no debe estar centrada tanto en el qué (simple conocimiento) como en el para qué, es decir, en valorar el grado en que las actitudes y valores promovidos han sido comprendidos por el alumnado, no para emitir una calificación moral, sino más bien para planificar y decidir las acciones educativas siguientes más adecuadas para el desarrollo del alumnado (Bolívar, 1995).

Las actitudes son tendencias o predisposiciones con componentes cognitivos, conductuales, pero sobre todo emotivos, positivos y negativos, hacia un determinado objeto de actitud. La conceptualización y la medición de actitudes se han desarrollado en el marco de la psicología social, desde hace más de setenta años (Morales,

Reboloso y Moya, 1994; Rodríguez, 1989; Stahlberg y Frey, 1990), con frecuentes polémicas sobre la validez de los instrumentos y los procesos de evaluación de las actitudes. La investigación de las actitudes relacionadas con la ciencia ha sido más tardía y ha discurrido de espaldas a la psicología social, por lo que ha estado condenada a repetir algunos de sus fallos. Así, diversas revisiones (Gardner, 1975a; Gauldy y Hukins, 1980; Schibeci, 1984; Shrigley y Koballa, 1992) han coincidido en señalar los defectos metodológicos claves observados en las investigaciones de evaluación de las actitudes en ciencias, atribuyéndolos a la falta de un marco teórico apropiado que, según los últimos autores, debería ser la psicología social.

El movimiento didáctico ciencia-tecnología-sociedad (CTS) tiene como uno de sus objetivos el desarrollo de las actitudes relacionadas con la ciencia en el alumnado y propone como referencia para su evaluación el cuerpo de conocimientos que emerge de los análisis históricos, filosóficos y sociológicos sobre la ciencia (Aikenhead, 1994a, 1994b; Bybee, 1987). En el espíritu de este movimiento está el deseo de ofrecer, a través de la educación de las actitudes relacionadas con la ciencia, una visión más auténtica de la ciencia y la tecnología en su contexto social, alejadas de imágenes mitificadas y sesgadas (cientifismo y tecnocracia) al mismo tiempo que se reconoce la tecnología, como una actividad diferente, integrada y equiparable con la ciencia, y no sólo como mera ciencia aplicada. La equiparación entre ciencia y tecnología plantea inmediatamente la naturaleza cargada de valores de las actividades científicas, de modo que la educación actitudinal –moral o ética– es una consecuencia inevitable de la educación CTS (Layton, 1994). Como afirma Ziman (1994), la debilidad de la ciencia tradicional no reside en lo que enseña sobre la naturaleza, sino en lo que no enseña, en particular, sus relaciones con la tecnología y la sociedad, vacío que pretende llenar la educación CTS.

Desde una mentalidad positivista, que cree en una ciencia neutral, objetiva, lógica, empírica, en la cual no queda lugar para otros factores distintos de la racionalidad y los hechos, y que es dominante en la investigación científica y entre el profesorado de ciencias (Abell y Smith, 1994; Blanco y Niaz, 1997; Mellado, 1998; Rubba y Harkness, 1993; Ruggieri, Tarsitani y Vicentini, 1993), es frecuente encontrar la conclusión de que las actitudes no tienen cabida en la educación en ciencias y, en consecuencia, tampoco se pueden evaluar, pues por su propio carácter (sesgadas, subjetivas, ilógicas, etc.) son opuestas a la naturaleza de la ciencia. Sin embargo, el conocimiento sobre la ciencia construido desde la reflexión filosófica y la investigación histórica y sociológica ha cuestionado ampliamente esta concepción y ha construido una visión más realista, diversa, dialéctica y cargada de valores de la ciencia, y estos rasgos, ineludiblemente, se deben proyectar en la educación y la evaluación. Así, como prueba de este carácter dialéctico, Alters (1997a) observa desacuerdos y diferencias sobre los principios básicos de la ciencia entre especialistas (filósofos de la ciencia), resultados que a su vez han sido cuestionados por la metodología empleada (Alters, 1997b;

Smith, Lederman, Bell, McComas y Clouh, 1997). Asimismo, Rubba, Schoneweg y Harkness (1996) utilizan cinco jueces para valorar diversas proposiciones de algunas cuestiones, informando desacuerdos entre jueces y que su escaso número produce dificultades en sus resultados, especialmente por las desviaciones provocadas por uno de los jueces. Sin embargo, a pesar del carácter dialéctico y multidimensional de las actitudes, la investigación coincide también en identificar algunas más ingenuas o inapropiadas (la credulidad experimentalista, el embelesamiento empiricista, el realismo ingenuo, la fe neutralista, el inductivismo, etc.), mientras otras actitudes han sido consideradas más adecuadas (el constructivismo social, la falibilidad de las teorías y las decisiones sociotécnicas, el carácter evolutivo de la ciencia, etc.).

En suma, las actitudes relacionadas con la ciencia no reflejan leyes demostradas empíricamente sino sistemas de valores, cuya evaluación no puede objetivarse ni calificarse con la misma sencillez que las respuestas a un test de química o de física (Lederman, 1986), y esta limitación apriorística es parte constitutiva de la actitud, y proyecta los principales problemas para su evaluación. El objetivo de la enseñanza de las actitudes relacionadas con la ciencia no debe ser sólo promover un punto de vista particular sobre la ciencia, sino abrir la formación a las distintas alternativas que aún son objeto de estudio, debate y, si cabe, elección, reconociendo la existencia de respuestas plurales a importantes cuestiones sobre la ciencia. De hecho, como muestran algunos estudios, la omisión de la enseñanza de las actitudes y valores de la ciencia, la falta de elección, están produciendo, por defecto, un tipo de profesorado y alumnado principalmente tradicional y positivista (Blanco y Niaz, 1997; Mellado, 1998).

Revisando toda esta compleja problemática conceptual e instrumental se ha propuesto una taxonomía multidimensional de las actitudes relacionadas con la ciencia (en plural, porque es básico admitir que existen muchos y diferentes objetos de actitud), que permite una definición más precisa y delimitada de cada objeto de actitud. Esta taxonomía propone cuatro dimensiones actitudinales básicas –actitudes hacia la ciencia escolar, la imagen social de la ciencia, las cuestiones específicas CTS y el conocimiento científico– y constituye un paso metodológico básico para una evaluación más válida y correcta de las actitudes y para la aceptación de las distintas actitudes como contenidos diferenciados e importantes de la educación en ciencias (Vázquez y Manassero, 1995, 1998a, 1998b).

LOS INSTRUMENTOS DE EVALUACIÓN DE LAS ACTITUDES RELACIONADAS CON LA CIENCIA

Este estudio está dirigido a concretar métodos de evaluación de las actitudes relacionadas con la ciencia, para facilitar su planificación y educación explícita en las áreas y materias de ciencias. Desde hace años, los clásicos

instrumentos de lápiz y papel empleados en la evaluación de las actitudes han sido repetidamente cuestionados por su escasa validez y fiabilidad, debido a la falta de precisión en la definición del objeto de actitud que se mide y a la ausencia de un constructo único y común a toda la escala, cuando no a la explícita multidimensionalidad del constructo, todas ellas diferentes formas de violación del supuesto de unidimensionalidad de constructo necesario para cualquier escala (Bratt, 1984; Munby, 1983; Zeidler, 1984). Dada la naturaleza dialéctica de las actitudes relacionadas con la ciencia, la ausencia de especificación de los supuestos filosóficos y de los modelos de la ciencia subyacentes en los cuestionarios invalidan y debilitan los datos e interpretaciones obtenidas mediante estos instrumentos (Aikenhead, 1988; Gardner, 1975b; 1996; Haladyna y Shaughnessy, 1982; Schibeci, 1984; Ormerod y Duckword, 1975; Shrigley y Koballa, 1992). Pero, sobre todos, el más grave defecto y también el más difícil de valorar y corregir es la inadecuación de criterio, es decir, la falta de ajuste o correspondencia entre el método o instrumento elegido para la medida y el objeto actitudinal a medir, es decir, entre lo que se quiere medir y lo que se mide realmente (Gauld y Hukins, 1980). Otros problemas señalados son la tendencia de los alumnos a responder para satisfacer las expectativas del profesor y la dificultad del lenguaje técnico para que investigadores y estudiantes perciban, entiendan e interpreten exactamente lo mismo cuando leen una frase de un cuestionario (doctrina de la percepción inmaculada). Todo ello es debido a los problemas semánticos de significado y comprensión de la terminología empleada, que pueden ser mitigados con una terminología más precisa o una buena explicación al alumnado. La ambigüedad del lenguaje fue estudiada por Aikenhead (1988) contrastando las respuestas escritas obtenidas de distintos instrumentos de evaluación actitudinal (Likert, párrafos escritos, elección múltiple derivada empíricamente y entrevistas semiestructuradas) con las respuestas en otras entrevistas complementarias. Este autor concluye que las entrevistas semiestructuradas reducen la ambigüedad casi a cero, pero exigen gran cantidad de tiempo para recoger y analizar los datos, de modo que los cuestionarios de elección múltiple derivados empíricamente son el método más operativo para reducir la ambigüedad.

En este marco se desarrolló el cuestionario VOSTS (Views on Science-Technology-Society), un conjunto de 114 cuestiones de opción múltiple construido a partir del análisis empírico de las respuestas escritas y las entrevistas de los estudiantes (empíricamente desarrollado) cuyo objetivo principal es superar las deficiencias metodológicas de los instrumentos tradicionales (Aikenhead y Ryan, 1989, 1992). La estructura conceptual del VOSTS abarca los objetos siguientes: *a*) definiciones de ciencia y tecnología; *b*) interacciones mutuas entre la ciencia, la tecnología y la sociedad; *c*) sociología externa de la ciencia, la sociología interna de la ciencia (características de científicos, construcción social de la tecnología y conocimiento científico); y *d*) naturaleza del conocimiento científico. Adaptando y refundiendo algunas de las cuestiones originales (Aikenhead y Ryan, 1992) y añadiendo otras cuestiones nuevas con el mismo forma-

Tabla I
Cuestión 10111 del COCTS junto con la categoría (adecuada, plausible o ingenua) asignada a cada opción por los jueces.

10111 Definir qué es la ciencia es difícil porque ésta es algo complejo y engloba muchas cosas. Pero la ciencia principalmente es:										
Para cada una de las frases siguientes, marca el número de la escala que represente mejor el grado de acuerdo entre tu propia opinión y la posición expuesta en la frase.	Grado de Acuerdo									CAT
	Bajo	Medio			Alto					
A. El estudio de campos tales como biología, química geología y física.	1	2	3	4	5	6	7	8	9	P
B. Un cuerpo de conocimientos, tales como principios, leyes y teorías que explican el mundo que nos rodea (materia, energía, y vida).	1	2	3	4	5	6	7	8	9	A
C. Explorar lo desconocido y descubrir cosas nuevas sobre el mundo y el universo y sobre cómo funcionan.	1	2	3	4	5	6	7	8	9	P
D. Realizar experimentos para resolver problemas de interés sobre el mundo que nos rodea.	1	2	3	4	5	6	7	8	9	P
E. Inventar o diseñar cosas (por ejemplo, corazones artificiales, ordenadores, vehículos espaciales).	1	2	3	4	5	6	7	8	9	I
F. Buscar y usar conocimientos para hacer de este mundo un lugar mejor para vivir (por ejemplo, curar enfermedades, solucionar la contaminación y mejorar la agricultura).	1	2	3	4	5	6	7	8	9	P
G. Una organización de personas (llamados científicos) que tiene ideas y técnicas para descubrir nuevos conocimientos.	1	2	3	4	5	6	7	8	9	P
H. Un proceso investigador sistemático y un conocimiento resultante.	1	2	3	4	5	6	7	8	9	A
I. No se puede definir la ciencia.	1	2	3	4	5	6	7	8	9	I

Si alguna de las frases siguientes es aplicable a las opciones anteriores, escribe la letra de la opción a su lado.

1. No lo entiendo.
2. No sé lo suficiente sobre este tema para elegir una opción.
3. Ninguna de estas opciones satisface básicamente mi opinión.

to, sugeridas por Rubba y Harkness (1993), se ha construido una versión española, COCTS (Cuestionario de Opiniones sobre Ciencia, Tecnología y Sociedad) con un centenar de cuestiones (Vázquez y Manassero, 1997, Manassero y Vázquez, 1998).

Las opciones de cada pregunta ofrecen un abanico plural de posibilidades actitudinales en cada cuestión a las que se aplica un formato de respuesta que consiste en seleccionar la opción que mejor se ajusta a la opinión del que responde (modelo de respuesta única). Este modelo resulta muy limitado metodológicamente, pues sólo permite comparaciones centradas en cada cuestión particular, pero no permite comparaciones test-retest o la verificación de hipótesis, procedimientos básicos de la estadística inferencial para conjuntos de cuestiones, más fiables que una sola cuestión. Para superar estas limitaciones, Rubba, Schoneweg y Harkness (1996) sugieren puntuar la respuesta única seleccionada en cada cuestión según una escala de valor previamente asignada a las proposiciones. De esta forma, las puntuaciones acumuladas en varias cuestiones pueden ser utilizadas en procedimientos de estadística inferencial. Aunque esta sugerencia constituye un avance, tiene importantes in-

convenientes de medida, por cuanto las puntuaciones obtenidas no discriminan respuestas muy diferentes, es decir, el método de medida no tiene la fidelidad necesaria para representar adecuadamente las actitudes evaluadas, pues una misma medida representa actitudes diferentes. Este inconveniente puede ser superado utilizando, para valorar cada una de las tres categorías de las respuestas, una métrica un poco más compleja y discriminadora que asigna 3,5 puntos a las categorías adecuadas, 1 punto a las categorías plausibles y 0 puntos a las categorías ingenuas (Vázquez y Manassero, 1999).

Sin embargo, el modelo de respuesta única tiene una limitación inherente imposible de ser superada por la elección de una métrica más fiel, pues no usa toda la información disponible en cada cuestión en todas las alternativas disponibles. Una única alternativa seleccionada sólo permite saber que la alternativa elegida es la más acorde con la actitud del que responde, pero no permite conocer nada de su opinión sobre las demás. Para superar este inconveniente, manteniendo las ventajas métricas (fidelidad de las medidas y discriminación de respuestas), es obvio que un modelo de respuesta múltiple resultaría más adecuado. Por ello, se ha sugeri-

Tabla II
Cuestión 10211 del COCTS junto con la categoría (adecuada, plausible o ingenua) asignada a cada opción por los jueces.

10211 Definir qué es la tecnología puede resultar difícil porque ésta sirve para muchas cosas. Pero la tecnología principalmente es:										
Para cada una de las frases siguientes, marca el número de la escala que represente mejor el grado de acuerdo entre tu propia opinión y la posición expuesta en la frase.	Grado de Acuerdo									CAT
	Bajo	Medio			Alto					
A. Muy parecida a la ciencia.	1	2	3	4	5	6	7	8	9	P
B. La aplicación de la ciencia.	1	2	3	4	5	6	7	8	9	I
C. Nuevos procesos, instrumentos, maquinaria, herramientas, aplicaciones, artilugios, ordenadores o aparatos prácticos para el uso de cada día.	1	2	3	4	5	6	7	8	9	P
D. Robots, electrónica, ordenadores, sistemas de comunicación, automatismos, máquinas.	1	2	3	4	5	6	7	8	9	P
E. Una técnica para construir cosas o una forma de resolver problemas prácticos.	1	2	3	4	5	6	7	8	9	P
F. Inventar, diseñar y probar cosas (por ejemplo, corazones artificiales, ordenadores y vehículos espaciales).	1	2	3	4	5	6	7	8	9	P
G. Ideas y técnicas para diseñar y hacer cosas; para organizar a los trabajadores, la gente de negocios y los consumidores; y para el progreso de la sociedad.	1	2	3	4	5	6	7	8	9	A
H. Saber cómo hacer cosas (por ejemplo, instrumentos, maquinaria, aparatos).	1	2	3	4	5	6	7	8	9	P

Si alguna de las frases siguientes es aplicable a las opciones anteriores, escribe la letra de la opción a su lado.

1. No lo entiendo.
2. No sé lo suficiente sobre este tema para elegir una opción.
3. Ninguna de estas opciones satisface básicamente mi opinión.

do un modelo de respuesta múltiple en la que cada persona valora el grado de acuerdo con cada una de las opciones presentes en la cuestión sobre una escala de 9 puntos. Este modelo de respuesta múltiple maximiza la información disponible en cada cuestión del COCTS y alcanza el mayor grado de precisión en la evaluación de las actitudes (Vázquez y Manassero, 1999).

La simple suma de las valoraciones asignadas (como hacen las escalas Likert) no es una buena medida de las actitudes, pues las distintas alternativas tienen diferentes valores actitudinales. Para interpretar las respuestas directas del modelo de respuesta múltiple como medidas actitudinales, se requiere la previa baremación de las distintas alternativas de cada cuestión y la propuesta de una métrica adecuada para las puntuaciones. La baremación está centrada en clasificar las alternativas en tres categorías:

Adecuadas (A): Si la frase expresa una opinión adecuada sobre el tema (coherente con los conocimientos de historia, epistemología y sociología de la ciencia).

Plausibles (P): Aunque no totalmente adecuada, la frase expresa algún aspecto adecuado.

Ingenuas (I): La frase expresa un punto de vista que no es ni adecuado ni plausible.

La baremación de las frases de cada cuestión mediante las categorías define un sistema local de significados y pesos que no sólo mejoran la eficiencia del COCTS (máxima información, medidas de alta fidelidad y posibilidad de estadística inferencial) sino que además evitan la objeción de la multidimensionalidad contra los instrumentos actitudinales, ya que todas las medidas se hacen sobre las tasaciones de cada alternativa dentro de la misma cuestión.

Para completar el modelo de respuesta múltiple se sugiere una métrica que permite obtener una valoración global y sintética de la actitud en cada cuestión a través del índice global actitudinal (rango -1, +1), construido sobre la base del carácter adecuado, plausible o ingenuo de cada una de las posiciones. Para las posiciones categorizadas como adecuadas, la actitud más valiosa (rechazable) será aquella que reconoce un ajuste o acuerdo total (bajo) con este tipo de posiciones asignando puntuaciones directas altas (bajas). Para las posiciones categorizadas como ingenuas, la actitud más valiosa (rechazable) será aquella que muestra un ajuste o acuerdo nulo (total) con este tipo de posiciones asignando puntuaciones directas bajas (altas). Para las posiciones plausibles, caracterizadas por una adecuación parcial de su contenido, las actitudes más valiosas (rechazables) serían aquellas que muestran un ajuste parcial o mediano (extremo) con estas posiciones expresado en puntuaciones inter-

Tabla III
Significado y asignaciones de puntos en la escala de valoración para el modelo de respuesta múltiple y los procedimientos de cálculo de los índices.

Categorías	Número de posiciones	Escala de valoración: significado de las puntuaciones									Cálculo de las puntuaciones directas			Cálculo de los índices de actitud de categoría		
		9	8	7	6	5	4	3	2	1	Máximo	Fórmula	Mínimo	Máx.	Índices	Mín.
Escala directa		Total	Casi total	Alto	Parcial alto	Parcial bajo	Bajo	Casi nulo	Nulo							
Adecuadas	N_a	4	3	2	1	0	-1	-2	-3	-4	$+4N_a$	$\sum a_j$	$-4N_a$	1	$I_a = \sum a_j / 4N_a$	-1
Plausibles	N_p	-2	-1	0	1	2	1	0	-1	-2	$+2N_p$	$\sum p_j$	$-2N_p$	1	$I_p = \sum p_j / 2N_p$	-1
Ingenuas	N_n	-4	-3	-2	-1	0	1	2	3	4	$+4N_n$	$\sum n_j$	$-4N_n$	1	$I_n = \sum n_j / 4N_n$	-1
Total	N										Índice de actitud global de cada ítem			1	$I = (I_a + I_p + I_n) / 3$	-1

a_j : Puntuación de valoración directa para la posición «adecuada» j.
 p_j : Puntuación de valoración directa para la posición «plausible» j.
 n_j : Puntuación de valoración directa para la posición «ingenua» j.
 Σ : Suma las puntuaciones directas desde $j = 1$ a $j = N_a$ ($j = N_p$ o $j = N_n$) para las categorías «adecuadas», «plausibles» o «ingenuas».

medias (extremas, altas o bajas). Para obtener el índice de actitud global del modelo de respuesta múltiple se transforman las puntuaciones directas sobre la escala de nueve puntos en puntuaciones finales de la actitud global según el algoritmo resumido en la tabla III.

El resultado, el índice de actitud global en la cuestión, es un número comprendido en el rango +1, -1, que tiene el mismo significado que los índices de actitud de categoría. Si el índice es positivo, la actitud es valiosa, tanto mejor cuanto más se acerca al valor de unidad. Si el índice es negativo, la actitud es ingenua, tanto más ingenua cuanto más se acerca al valor de unidad negativa (Vázquez y Manassero, 1999). El índice de actitud global es independiente de la cuestión a responder, del número de alternativas o posiciones que tiene y del número de posiciones adecuadas, plausibles e ingenuas que posee, de modo que es un parámetro invariante, puesto que es independiente de la cuestión y de la estructura que posee ésta. Esta métrica permite obtener un índice total semejante, con el mismo significado, cuando se aplica un conjunto de cuestiones, sumando los índices de actitud global en cada cuestión y dividiendo por el número de cuestiones aplicadas.

BAREMACIÓN DE LAS ALTERNATIVAS MEDIANTE JUECES

Las cuestiones fueron sometidas a estudio de un panel de once jueces, expertos en cuestiones CTS (investigadores en CTS o profesores de epistemología o sociología), a

quienes se pidió que valoraran el grado de adecuación de cada una de las proposiciones de cada cuestión, etiquetadas con una letra mayúscula, desde la perspectiva de los conocimientos de historia, filosofía y sociología de la ciencia. Para cuantificar esta valoración se propuso tasar cada una de ellas sobre una escala de nueve puntos, cuyo significado es el siguiente:

Adecuada (A): La proposición expresa una opinión apropiada desde la perspectiva de los conocimientos de historia, filosofía y sociología de la ciencia (7, 8, 9 puntos).

Plausible (P): Aunque no completamente adecuada, la proposición expresa algunos aspectos apropiados, desde la perspectiva de los conocimientos de historia, filosofía y sociología de la ciencia (4, 5, 6 puntos).

Ingenua (I): La proposición expresa una opinión inapropiada o no plausible (1, 2, 3 puntos).

Las valoraciones de los jueces permitieron asignar cada frase a una de las tres categorías utilizando dos criterios principales, la puntuación media aritmética de todos los jueces y las mayorías en favor de una u otra categoría de los jueces. También se tuvieron en cuenta algunos mecanismos correctores como, por ejemplo, evitar la tendencia a la centralización, que favorecería en exceso la categoría de plausibles, ya que, al ocupar el centro de la escala numérica, recoge las desviaciones por arriba y por abajo. Las categorías de cada una de las opciones aparecen señaladas en las tablas de las cuestiones en la columna de la derecha con las iniciales correspondientes (A, P, I).

INVESTIGACIÓN DIDÁCTICA

Tabla IV
Cuestión 10412 del COCTS junto con la categoría (adecuada, plausible o ingenua) asignada a cada opción por los jueces.

10412 <i>¿La ciencia influye en la tecnología?</i>										
Para cada una de las frases siguientes, marca el número de la escala que represente mejor el grado de acuerdo entre tu propia opinión y la posición expuesta en la frase.	Grado de Acuerdo									CAT
	Bajo	Medio			Alto					
A. La ciencia no influye demasiado en la tecnología.	1	2	3	4	5	6	7	8	9	I
B. Tecnología es ciencia aplicada.	1	2	3	4	5	6	7	8	9	I
C. El avance en ciencia conduce a nuevas tecnologías.	1	2	3	4	5	6	7	8	9	P
D. La ciencia se hace más valiosa cuando se usa en tecnología.	1	2	3	4	5	6	7	8	9	P
E. La ciencia es el conocimiento base para la tecnología.	1	2	3	4	5	6	7	8	9	P
F. Los conocimientos de la investigación científica aplicada se usan más en tecnología que los conocimientos de la investigación científica pura.	1	2	3	4	5	6	7	8	9	P
G. La tecnología es la aplicación de la ciencia para mejorar la vida.	1	2	3	4	5	6	7	8	9	I
<p><i>Si alguna de las frases siguientes es aplicable a las opciones anteriores, escribe la letra de la opción a su lado.</i></p> <p>1. No lo entiendo. 2. No sé lo suficiente sobre este tema para elegir una opción. 3. Ninguna de estas opciones satisface básicamente mi opinión.</p>										

Tabla V
Cuestión 10413 del COCTS junto con la categoría (adecuada, plausible o ingenua) asignada a cada opción por los jueces.

10413 <i>¿La tecnología influye en la ciencia?</i>										
Para cada una de las frases siguientes, marca el número de la escala que represente mejor el grado de acuerdo entre tu propia opinión y la posición expuesta en la frase.	Grado de Acuerdo									CAT
	Bajo	Medio			Alto					
A. La tecnología no influye en gran medida sobre la ciencia.	1	2	3	4	5	6	7	8	9	I
B. La capacidad para crear tecnología marca el valor del conocimiento científico.	1	2	3	4	5	6	7	8	9	P
C. La disponibilidad de tecnología influye en la dirección de la investigación científica.	1	2	3	4	5	6	7	8	9	A
D. Los avances tecnológicos conducen a progresos en la ciencia.	1	2	3	4	5	6	7	8	9	A
E. La tecnología se usa por la sociedad para descubrir nuevos conocimientos científicos.	1	2	3	4	5	6	7	8	9	P
F. La tecnología suministra herramientas y técnicas para la ciencia.	1	2	3	4	5	6	7	8	9	A
G. La tecnología es la aplicación de la ciencia para mejorar la vida.	1	2	3	4	5	6	7	8	9	I
<p><i>Si alguna de las frases siguientes es aplicable a las opciones anteriores, escribe la letra de la opción a su lado.</i></p> <p>1. No lo entiendo. 2. No sé lo suficiente sobre este tema para elegir una opción. 3. Ninguna de estas opciones satisface básicamente mi opinión.</p>										

Una diferencia fundamental del COCTS respecto a otros cuestionarios es la pluralidad de todas sus cuestiones, puesto que cada una de las alternativas recoge múltiples perspectivas epistemológicas. Por decirlo más sencillamente, el problema de los cuestionarios clásicos, por ejemplo, es que cada una de sus frases reflejaba una filosofía (que se pretendía que fuera la adecuada), de modo que la evaluación estaba en relación con esa filosofía subyacente (y ninguna otra). En el caso del COCTS, las alternativas definen un espectro plural que recoge posiciones de diversas filosofías y con diversos grados, de modo que el índice actitudinal marca la posición del encuestado en relación con toda la pluralidad epistemológica existente sobre un tema.

UNA APLICACIÓN EMPÍRICA A LA MEDIDA DE ACTITUDES

Como ejemplificación del modelo de respuesta múltiple se han aplicado algunas cuestiones del COCTS, previamente categorizadas por un conjunto de jueces (Tablas I, II, IV y V) estudiantes de bachillerato (17-18 años), cuyas respuestas numéricas se han valorado con la métrica expuesta para evaluar actitudes. Por razones de espacio se consideran sólo cuatro cuestiones referidas a la definición de ciencia (10111), definición de tecnología (10211), influencia de la ciencia en la tecnología (10412) e influencia de la tecnología en la ciencia (10413).

Aplicando la métrica explicitada anteriormente se calculan los índices globales actitudinales para cada cuestión como promedio entre los tres índices referidos a cada una de las tres categorías. Las actitudes obtenidas son positivas en las cuatro cuestiones, aunque con diferencias importantes entre ellas, pues mientras que la definición de la ciencia y la influencia de la tecnología sobre la ciencia tienen índices actitudinales moderadamente positivos, las otras dos cuestiones referidas a la definición de tecnología y la influencia de la ciencia sobre la tecnología muestran actitudes nulas (Tabla VI). Las actitudes moderadamente adecuadas en las dos primeras cuestiones indican ya una formación de los estudiantes alejada de cualquier estándar deseable, que es todavía más bajo en el caso de las dos últimas cuestiones cuyas actitudes claramente indefinidas son un indicador claro de una falta absoluta de formación. Ni es el objeto de este estudio, ni se dispone de espacio para profundizar en la descripción pormenorizada de las actitudes específicas de los estudiantes sobre estas cuestiones que se han abordado con más profundidad en otro lugar (Manassero y Vázquez, 1998), pero la potencia de la metodología diagnóstica de actitudes que se propugna puede profundizar hasta el detalle de la contribución de cada una de las alternativas de cada cuestión para configurar la actitud global. Por ejemplo, la línea «Media-índices» de la tabla VII evidencia que las alternativas E (inventar o diseñar cosas, por ejemplo, corazones artificiales, ordenadores, vehículos espaciales), F (buscar y usar conocimientos para hacer de este mundo un lugar mejor para vivir, por ejemplo, curar enfermedades, solucionar la contaminación y mejorar la agricultura), G (una organi-

zación de personas que tienen ideas y técnicas para descubrir nuevos conocimientos) y H (un proceso investigador sistemático y el conocimiento resultante) exhiben las actitudes más deficientes de los estudiantes en la primera cuestión evaluada (10111).

Al lado del índice actitudinal medio se ha calculado también la media de todas las alternativas de cada cuestión, observándose que sus valores no se desvían ostensiblemente de los valores calculados mediante el método del párrafo anterior para el índice de actitud global. Sólo en el caso de la segunda cuestión, la definición de *tecnología*, el índice actitudinal es claramente inferior al índice promediado sobre todas las alternativas.

Analizando los índices actitudinales en cada una de las categorías para cada cuestión (ingenuas, plausibles y adecuadas), se puede observar cómo la contribución a la actitud de las categorías adecuadas y plausibles es muy similar en todas las cuestiones, en tanto que los índices correspondientes a las categorías ingenuas tienden a dar índices más bajos que todos los demás (con excepción de la primera cuestión 10111), diferencia que es muy notoria en el caso de la cuestión referida a la definición de la tecnología.

Tabla VI
Índices actitudinales de las cuatro cuestiones evaluadas por categorías e índices de actitud global. (Media-categorías: promedio directo de los índices de todas las alternativas de una cuestión).

Índices	C10111	C10211	C10412	C10413
Media-adecuadas	0,125	0,131	0,000	0,234
Media-plausibles	0,171	0,139	0,095	0,155
Media-ingenuas	0,315	-0,262	-0,004	0,048
Media todas alternativas	0,167	0,102	0,073	0,148
Índice actitudinal global	0,204	0,003	0,046	0,146

El perfil de respuesta del alumnado en una cuestión puede compararse gráficamente respecto a los índices medios de respuesta de la muestra total en cada alternativa, lo cual permite identificar las alternativas en las que el alumnado tiene las actitudes mejores, peores e intermedias con relación a la muestra. La tabla VII, en las últimas filas, recoge el perfil de respuesta de la cuestión primera, definición de *ciencia*, observándose cómo las alternativas A, B, C, D e I tienen las actitudes más positivas, mientras que las alternativas F y H tienen los índices más bajos. Con estos parámetros, el perfil de respuesta ofrece un patrón de diagnóstico claro sobre la actitudes de los estudiantes, evidenciando los puntos fuertes y los puntos débiles en cada tema. Estos resultados pueden utilizarse como una guía diagnóstica de ideas previas para afrontar la enseñanza de estas cuestiones en la práctica.

Análogamente, las puntuaciones individualizadas, invariantes y equivalentes permiten realizar comparaciones entre grupos, por ejemplo, entre hombres y mujeres o entre un grupo de control y otro grupo experimental, aplicaciones al análisis de diferencias test-retest y, en general, todas las pruebas de comparación estadística propias de la estadística inferencial para el contraste de hipótesis. La tabla VII, por ejemplo, en las dos últimas líneas contiene los perfiles de chicos y chicas, observándose cómo las chicas tienen actitudes más positivas que los chicos en las alternativas B, C y D, mientras que los chicos tienen actitudes más positivas que las chicas en las alternativas E y F. Asimismo, refiriéndonos a los índices actitudinales globales de la tabla VI, se podrían aplicar pruebas de significación estadística para establecer el grado de significación de las diferencias entre grupos en una investigación para determinar la influencia de un programa o un currículo en la actitudes, o las diferencias o ganancias producidas entre grupos diferentes, etc.

Finalmente, en el caso de aplicar un conjunto de cuestiones y desear obtener una puntuación globalizada como representación totalizadora de la actitud global, el carácter homogéneo e invariante de los índices globales definidos en cada cuestión puede permitir definir puntuaciones totales de actitud. Por ejemplo, en la última línea de la tabla VI se han calculado los parámetros correspondientes de la tabla para las cuatro cuestiones aplicadas como promedio de los parámetros obtenidos en cada cuestión; sobre estos parámetros totales se pueden seguir aplicando todas las comparaciones y los tests de la estadística inferencial.

Comparación con el modelo de respuesta única

Dentro del modelo de respuesta múltiple, el modelo de respuesta única es sólo un caso límite en el cual sólo existe la respuesta que tiene la máxima puntuación o grado de acuerdo del encuestado. Como resulta evidente, toda la información referente a las otras alternativas que no son ésta, no se tiene en cuenta y es una información que se pierde en el modelo de respuesta única y es la base principal de la superioridad del modelo de respuesta múltiple. Pero además existe otra desventaja del modelo de respuesta única que ya se había anticipado (Vázquez y Manassero, 1999) y es la posible igualdad en la preferencia entre dos (o más alternativas), de manera que la obligación de elegir una de ellas en detrimento de las otras (en el modelo de respuesta única), supone una nueva forma de limitar y deformar la respuesta que no sucede en el modelo de respuesta múltiple.

La comparación empírica entre ambos modelos confirma estos rasgos y evidencia la mayor precisión del modelo de respuesta múltiple en la evaluación de las actitudes. Para realizar la comparación se considera que la alternativa que alcanza la mayor puntuación del modelo de respuesta múltiple sería la opción seleccionada en el modelo de respuesta única. En la tabla VII se muestran las respuestas directas del alumnado a la cuestión de definición de la ciencia (10111), que se ha

tomado como ejemplo. En las respuestas de la tabla se observa que aproximadamente la mitad del alumnado (9) está en la situación de asignar su puntuación máxima a más de una opción, es decir, que existen numerosas valoraciones cuya máxima puntuación no es única, puesto que más de una opción reciben la misma puntuación máxima. Dos de los estudiantes tienen una multiplicidad de valoración máxima que afecta a cuatro opciones, otros dos tienen tres opciones con la valoración máxima y en otros cinco coinciden dos opciones con la valoración máxima. En estos casos de respuesta máxima sobre varias opciones simultáneamente, el modelo de respuesta única introduce en la evaluación actitudinal una imprecisión adicional, pues la exigencia de respuesta única obliga, al que responde, a elegir una sola de sus múltiples opciones máxima e igualmente preferidas y a despreciar las otras igualmente valoradas, y por ello igualmente válidas, lo cual supone una cierta deformación de la opinión de quien responde y, como ya se había notado antes, una pérdida de información valiosa para la definición de la actitud.

Finalmente, contabilizando todas las respuesta máximas obtenidas para calcular la distribución porcentual de respuestas únicas (Tabla VII, porcentajes de puntuaciones máximas) y comparándola con la distribución de porcentajes de una muestra más grande (Tabla VII, porcentajes de una muestra amplia) pero de características similares a ésta en edad y estudios (Vázquez y Manassero, 1997), se observan discrepancias importantes en algunas opciones. Aunque no se puede concluir que el modelo de respuesta única sea la causa única de estas discrepancias –debido a la falta de representatividad de la muestra de este estudio–, sí sería esperable que estas diferencias existieran, debido al gran número de respuestas máximas que el modelo de respuesta única desprecia cuando existe una multiplicidad de respuestas máximas, y que son decididas por los estudiantes con criterios para seleccionar la respuesta única que no se puede conocer. En suma, empíricamente, el modelo de respuesta múltiple permite un conocimiento más exacto y mucho más informado de las actitudes de los estudiantes que el modelo de respuesta única.

DISCUSIÓN

Las cuestiones COCTS se han convertido en el banco de ítems mejor fundado y más empleado para evaluar actitudes, por su diseño empíricamente desarrollado que le hace acreedor de una buena validez. Sin embargo, su forma de respuesta original, en la que el estudiante sólo elige una de las alternativas, la que cuadra más con su opinión, tiene dos limitaciones principales: *a)* carece de un sistema de puntuación que permita emplear sus resultados en procedimientos de estadística inferencial; y *b)* pierde toda la información relativa a las opciones no seleccionadas, que incluye la existencia de otras opciones equipotentes con la única seleccionada.

En otro estudio (Vázquez y Manassero, 1999) se han analizado las insuficiencias métricas del procedimiento

INVESTIGACIÓN DIDÁCTICA

Tabla VII

Respuestas directas de los estudiantes a la cuestión 10111 (definición de ciencia) y distribución de respuestas sobre cada una de las alternativas.

OPCIONES	10111 A	10111 B	10111 C	10111 D	10111 E	10111 F	10111 G	10111 H	10111 I
CATEGORÍA ASIGNADA	Plausible	Adecuada	Plausible	Plausible	Ingenua	Plausible	Plausible	Adecuada	Ingenua
Sujeto 1	7	7	7	3	2	4	4	4	1
Sujeto 2	3	3	5	2	5	3	7	6	5
Sujeto 3	5	5	5	4	4	4	5	4	2
Sujeto 4	8	8	9	5	1	3	1	3	3
Sujeto 5	5	6	5	5	6	6	4	5	8
Sujeto 6	6	8	4	5	6	6	8	7	4
Sujeto 7	2	3	8	8	3	8	2	2	4
Sujeto 8	1	5	4	3	5	7	5	2	5
Sujeto 9	6	8	4	7	6	2	1	5	2
Sujeto 10	5	7	6	8	8	9	9	5	4
Sujeto 11	8	9	7	7	8	9	8	7	2
Sujeto 12	5	7	8	4	4	5	6	A	1
Sujeto 13	5	5	7	8	4	8	3	6	3
Sujeto 14	7	4	4	5	4	5	2	4	7
Sujeto 15	3	6	7	7	3	7	3	7	C
Sujeto 16	7	8	5	5	4	7	4	7	5
Sujeto 17	7	5	6	5	6	8	6	A	7
Sujeto 18	7	5	6	6	7	9	6	A	6
Sujeto 19	C	3	3	4	1	7	B	A	A
Sujeto 20	6	8	4	5	6	6	5	7	2
Sujeto 21	6	7	8	7	5	7	4	3	1
Frecuencias puntuaciones máximas	3	7	7	3	0	8	4	1	2
Porcentajes de puntuaciones máximas	8	19	19	8	0	25	11	3	6
Porcentajes de una muestra amplia	18	32	15	3	0	10	3	11	5
Media-índices	0,23	0,26	0,26	0,33	0,08	-0,02	0,08	-0,01	0,30
Desv-típica	0,60	0,46	0,56	0,58	0,50	0,60	0,67	0,45	0,54
Media-hombres	0,25	0,00	0,14	0,07	0,39	0,14	0,08	0,00	0,40
Media- mujeres	0,21	0,39	0,32	0,46	-0,07	-0,11	0,07	-0,02	0,27

sugerido por Rubba y otros (1996) para puntuar la única respuesta seleccionada y se ha propuesto un nuevo sistema que evite esos inconvenientes. Además, para aprovechar la gran cantidad de información contenida en todas las alternativas de las cuestiones COCTS y superar las insuficiencias inherentes al modelo de respuesta única se sugirió un modelo de respuesta múltiple, en el que todas las alternativas del ítem son valoradas por los encuestados. Una métrica de puntuación global integra las respuestas múltiples emitidas en una valoración única de la actitud global, en relación con su adecuación respecto a los conocimientos de historia, epistemología y sociología de la ciencia.

Este modelo se podría considerar como una nueva versión de los clásicos cuestionarios Likert de actitudes hacia la ciencia, formados por una colección de frases cuyas puntuaciones «de acuerdo» y «desacuerdo» son sumadas para dar una puntuación total y aquejados de problemas metodológicos como la falta de unidimensionalidad de la escala, reflejo de un constructo único al cual hipotéticamente deben ajustarse todas las frases, supuesto que se cumple muy raramente, porque el constructo no se ha definido bien o porque el constructo está formado por diversos componentes muy diferentes (multidimensional), de modo que la validez métrica de las puntuaciones y las consecuencias que se derivan de ellas

son débiles (Gardner, 1996). Sin embargo, estas objeciones no son aplicables al modelo propuesto aquí, porque el constructo se ha definido en cada cuestión y porque las puntuaciones se limitan a cada cuestión. El constructo medido por el sistema de puntuación definido para las respuestas múltiples es la actitud global respecto a una concepción adecuada de la ciencia, conformada por la historia, la epistemología y la sociología de la ciencia. Esta actitud se mide independiente e individualmente en cada uno de los temas planteados, es decir, en cada cuestión. Este procedimiento garantiza la unidimensionalidad del constructo valorado para cada cuestión, pues todas se refieren al mismo tema específico de contenido.

Por otro lado, las escalas Likert solicitan el acuerdo o desacuerdo con las proposiciones cuya asignación origina directamente la puntuación correspondiente. En el caso del modelo de respuesta múltiple, la situación es diferente: la puntuación sobre el ajuste de una posición determinada es más diversificada, a través de un sistema más complejo de valoración, según una tripleta de categorías asignadas a cada posición por un panel de jueces. La contribución de cada posición a la actitud global es, por este método, más precisa y matizada, y el resultado sobre la adecuación no es tan grosero como el resultado de tratar todas las frases, simplemente, como positivas o negativas. Los resultados empíricos de este estudio demuestran la utilidad del nuevo método de valoración del COCTS para la evaluación general de actitudes, junto con las ventajas y detalles diferenciales que potencian el instrumento como un elemento esencial.

Con independencia de las objeciones que se puedan aducir contra el modelo de respuesta múltiple, su calidad principal radica en que la actitud global se construye sobre la base de la información contenida en todas las opciones y no sólo con la información procedente de una de ellas, la que se considera más adecuada en el modelo de respuesta única. Las ventajas de este procedimiento no necesitan argumentarse, porque son obvias por sí mismas (varios informan más que uno), y además tiene la ventaja que el modelo de respuesta única se deduce de una manera natural del modelo de respuesta múltiple,

como un caso límite (conservando sólo la posición que obtiene la valoración más alta). Incluso, en este caso, el modelo de respuesta múltiple mejora la precisión de la actitud respecto al modelo de respuesta única, conociendo la valoración asignada a la posición y la posible existencia de otras opciones coincidentes en intensidad con la única. Como se demuestra en este estudio, esta última posibilidad no sólo no es irrelevante sino que es muy frecuente, pues se ha comprobado que casi la mitad de los que responden seleccionan más de una opción con preferencia máxima, los cuales muestran un cierto grado de equipotencia entre diferentes opciones, rasgo ya identificado en otros estudios (Acevedo, 2000; Manassero y Vázquez, 1998). La actitud global construida con la información de todas las opciones está más fundamentada, más informada y es más precisa que las anteriores.

La aplicación empírica presentada permite apreciar el valor de las metodologías empleadas en la evaluación de las actitudes, que potencian extraordinariamente los usos y aplicaciones del COCTS. Estas cuestiones pueden ser utilizadas en investigación en didáctica de la ciencia, haciendo evaluaciones globales, comprobando hipótesis inferidas estadísticamente, realizando comparaciones estadísticas entre grupos, etc. Pero también puede ser aplicado directamente en el aula, no sólo como instrumento de diagnóstico de actitudes, sino incluso como guía curricular para las actitudes relacionadas con la ciencia en la enseñanza de las ciencias, pues el banco de ítems COCTS tiene desarrolladas todas las dimensiones y los temas sobre los cuales se quieren plantear objetivos educativos actitudinales y que se han sistematizado en una taxonomía de las actitudes relacionadas con la ciencia y la tecnología (Vázquez y Manassero, 1997).

AGRADECIMIENTOS

A todos los expertos cuya participación como jueces de las cuestiones COCTS han permitido clasificarlas en algunas de las tres categorías y a un revisor por sus fructíferas sugerencias.

REFERENCIAS BIBLIOGRÁFICAS

- ABELL, S.K. y SMITH, D.C. (1994). What is science?: preservice elementary teachers' conception of the nature of science. *International Journal of Science Education*, 16(4), pp. 475-487.
- ACEVEDO, J.A. (2000). Algunas creencias sobre el conocimiento científico de los profesores de secundaria en formación inicial. *Bordón*, 52(1), pp. 5-16.
- AIKENHEAD, G.S. (1987). High School Graduates' Beliefs About Science-Technology-Society. III. Characteristics and Limitations of Scientific Knowledge. *Science Education*, 71(4), pp. 459-487.
- AIKENHEAD, G.S. (1988) An analysis of four ways of assessing student beliefs about STS topics. *Journal of Research in Science Teaching*, 25, pp. 607-627.
- AIKENHEAD, G. (1994a). Consequences to learning science through STS: a research perspective, en Solomon, J. y Aikenhead, G. (eds.). *STS education: International perspectives on reform*, pp. 169-186. Nueva York: Teachers College Press.
- AIKENHEAD, G. (1994b). What is STS science teaching?, en Solomon, J. y Aikenhead, G. (eds.). *STS education: International Perspectives on Reform*, pp. 47-59. Nueva York: Teachers College Press.

- AIKENHEAD, G.S., FLEMING, R.G. y RYAN, A.G. (1987). High School Graduates' Beliefs About Science-Technology-Society. I. Methods and Issues in Monitoring Students Views. *Science Education*, 71(2), pp. 145-161.
- AIKENHEAD, G.S. y RYAN, A.G. (1989). The development of a multiple choice instrument for monitoring views on Science-Technology-Society topics. Final Report of SSHRCC Grant: Author.
- AIKENHEAD, G.S. y RYAN, A.G. (1992). The development of a new instrument: «Views on Science-Technology-Society» (VOSTS). *Science Education*, 76(5), pp. 477-492.
- ALONSO, M., GIL, D. y MARTÍNEZ, J. (1992). Los exámenes de física en la enseñanza por transmisión y en la enseñanza por investigación. *Enseñanza de las Ciencias*, 10, pp. 127-138.
- ALTERS, B.J. (1997a). Whose nature of science? *Journal of Research in Science Teaching*, 34(1), pp. 39-55.
- ALTERS, B.J. (1997b). Nature of science: a diversity or uniformity of ideas? *Journal of Research in Science Teaching*, 34(10), pp. 1105-1108.
- ÁLVAREZ, J.M. (1985). *Didáctica, currículo y evaluación. Ensayos sobre cuestiones didácticas*. Barcelona: Alamex.
- BLANCO, R. y NIAZ, M. (1997). Epistemological beliefs of students and teachers about the nature of science: from a «Baconian inductive ascent» to the «irrelevance» of scientific laws. *Instructional Science*, 25, pp. 203-231.
- BOLÍVAR, A. (1995). *La evaluación de las actitudes*. Madrid: Anaya.
- BRATT, M. (1984). Further Comments on the Validity Studies of Attitude Measures in Science Education. *Journal of Research in Science Teaching*, 21(9), p. 951.
- BYBEE, R.W. (1987). Science education and the Science-Technology-Society (STS) theme. *Science Education*, 71(5), pp. 667-683.
- CAMPS, V. (1994). *Los valores de la educación*. Madrid: Anaya.
- GARDNER, P.L. (1975a). Attitudes to science: A review. *Studies in Science Education*, 2, pp. 1-41.
- GARDNER, P.L. (1975b). Attitude measurement: A critique of some recent research. *Education Research*, 17, pp. 101-105.
- GARDNER, P.L. (1996). The dimensionality of attitude scales: a widely misunderstood idea. *International Journal of Science Education*, 18, pp. 913-919.
- GAULD, C.F. y HUKINS, A.A. (1980). Scientific attitudes: A review. *Studies in Science Education*, 7, pp. 129-161.
- HALADYNA, T. y SHAUGHNESSY, J. (1982). Attitudes towards science: A quantitative synthesis. *Science Education*, 66, pp. 547-563.
- KEMPA, R. (1985). *Assessment in science*. Cambridge: Cambridge University Press.
- LAYTON, D. (1994). STS in the school curriculum: A movement overtaken by history?, en Solomon, J. y Aikenhead, G. (eds.). *STS education: International perspectives on reform*, pp. 32-44. Nueva York: Teachers College Press.
- LEDERMAN, N. (1986). Students' and teachers' understanding of the nature of science: A reassessment. *School Science and Mathematics*, 86(2), pp. 91-99.
- MANASSERO, M.A. y VÁZQUEZ, A. (1998). *Opiniones sobre ciencia, tecnología i societat*. Palma de Mallorca: Conselleria d'Educació, Cultura i Esports.
- MELLADO, V. (1998). Preservice teachers' classroom practices and their conceptions of the nature of science, en Fraser, B.J. y Tobin, K.G. (eds.). *International Handbook of Science Education*, pp. 1093-1110. Londres: Kluwer Academic Publishers.
- MORALES, J.F., REBOLLOSO, E. y MOYA, M. (1994). Actitudes, en Morales, J.F., Moya, M. Reboloso, E. Fernández, J.M., Huici, C., Marqués, J., Páez, D. y Pérez, J.A. *Psicología social*, pp. 495-566. Madrid: Mcgraw-Hill.
- MUNBY, H. (1983). Thirty studies involving the «Scientific Attitude Inventory»: What confidence can we have in this instrument? *Journal of Research in Science Teaching*, 20, 141-162.
- ORMEROD, M.B. y DUCKWORTH, D. (1975). *Pupils attitudes' to science: a review of research*. Windsor: NFER Publishing Co.
- RODRÍGUEZ, A. (1989). Interpretación de las actitudes, en Rodríguez, A. y Seoane, J. (coord.). *Creencias, actitudes y valores*, pp. 199-314. Madrid: Alhambra.
- RODRÍGUEZ, L.M., GUTIÉRREZ, F.A. y MOLLEDO, J. (1992). Una propuesta integral de evaluación en ciencias. *Enseñanza de las Ciencias*, 10, pp. 254-267.
- RUBBA, P.A., SCHONEWEG, C. y HARKNESS, W.L. (1996). A new scoring procedure for the Views on Science-Technology-Society instrument. *International Journal of Science Education*, 18, pp. 387-400.
- RUBBA, P.A. y HARKNESS, W.L. (1993). Examination of preservice and in-service secondary science teachers' beliefs about Science-Technology-Society interactions. *Science Education*, 77, pp. 407-431.
- RUGGIERI, R., TARSITANI, C. y VICENTINI, M. (1993). The images of science of teachers in Latin countries. *International Journal of Science Education*, 15(4), pp. 383-393.
- RYAN, A.G. y AIKENHEAD, G.S. (1992). Students' preconceptions about epistemology of science. *Science Education*, 76(6), pp. 559-580.
- RYAN, A.G. (1987). High School Graduates' Beliefs About Science-Technology-Society. IV. The Characteristics of Scientists. *Science Education*, 71(4), pp. 489-510.
- SATTERLY, D. y SWANN, N. (1988). Los exámenes referidos al criterio y al concepto en ciencias. Un nuevo sistema de evaluación. *Enseñanza de las Ciencias*, 6, pp. 278-284.
- SCHIBECI, R.A. (1984). Attitudes to science: Un update. *Studies in Science Education*, 11, pp. 26-59.
- SHRIGLEY, R.L. y KOBALLA Jr., T.R. (1992). A decade of attitude research based on Hovland's learning model. *Science Education*, 76(1), pp. 17-42.
- SMITH, M.U., LEDERMAN, N.G., BELL, R.L., McCOMAS, W.F. y CLOUGH, M.P. (1997). How great is disagreement about the nature of science: A response to Alters. *Journal of Research in Science Teaching*, 34(10), pp. 1101-1103.
- STAHLBERG, D. y FREY, D. (1990). Actitudes I: estructura, medida y funciones, en Hewstone, M., Stroebe, W., Codol, J.P. y Stephenson, G.M. (dir.). *Introducción a la psicología social*, pp. 149-170. Barcelona: Ariel.

- STUFFLEBEAM, D.L. y SHINKFIELD, A.J. (1987). *Evaluación sistemática. Guía teórica y práctica*. Madrid: Paidós-MEC.
- TENBRINK, T.D. (1983). *Evaluación. Guía práctica para profesores*. Madrid: Narcea.
- VÁZQUEZ, A., ACEVEDO, J.A. y MANASSERO, M.A. (2000). Progresos en la evaluación de actitudes relacionadas con la ciencia mediante el cuestionario de opiniones CTS, en Martins, I.P. (org.). *O Movimento CTS na Península Ibérica*, pp. 219-230. Aveiro: Universidade de Aveiro.
- VÁZQUEZ, A. y MANASSERO, M.A. (1995). Actitudes relacionadas con la ciencia: una revisión conceptual. *Enseñanza de las Ciencias*, 13(3), pp. 337-346.
- VÁZQUEZ, A. y MANASSERO, M.A. (1997). *Actitudes y valores relacionados con la ciencia, la tecnología y la sociedad en alumnado y profesorado. Implicaciones para la educación de las actitudes*. Memoria final de investigación. Madrid: MEC-CIDE.
- VÁZQUEZ, A. y MANASSERO, M.A. (1998a). *Actituds de l'alumnat relacionades amb la ciència, la tecnologia i la societat*. Palma de Mallorca: Conselleria d'Educació, Cultura i Esports.
- VÁZQUEZ, A. y MANASSERO, M.A. (1998b). Una propuesta de modelo integrado de aprendizaje como cambio conceptual, metodológico y actitudinal, en Banet, E. y de Pro, A. (coord.). *Investigación e innovación en la enseñanza de las ciencias*, V(1), pp. 148-158. Murcia: Universidad de Murcia.
- VÁZQUEZ, A. y MANASSERO, M.A. (1999). Response and scoring models for the «Views on Science.Technology-Society» instrument. *International Journal of Science Education*, 21(3), pp. 231-247.
- WALBERG, H.J. y HAERTEL, G.D. (eds.) (1990). *The International Encyclopedia of Educational Evaluation*. Oxford: Pergamon Press.
- ZEIDLER, D.L. (1984). Thirty studies involving the «Scientific Attitude Inventory»: what confidence can we have in this instrument. *Journal of Research in Science Teaching*, 21(3), pp. 341-342.
- ZIMAN, J. (1994). The rationale of STS. Education is in the approach, en Solomon, J. y Aikenhead, G. (eds.). *STS education: International perspectives on reform*, pp. 21-31. Nueva York: Teachers College Press.

[Artículo recibido en enero de 2000 y aceptado en octubre de 2000.]