

AVANCES METODOLÓGICOS EN LA INVESTIGACIÓN SOBRE
EVALUACIÓN DE ACTITUDES Y CREENCIAS CTS

José Antonio Acevedo Díaz
Inspección de Educación, Delegación Provincial de Educación y Ciencia de Huelva
Pilar Acevedo Romero
Departamento de Química Analítica. Universidad de Sevilla
María Antonia Manassero Mas
Departamento de Psicología, Universidad de las Islas Baleares
Ángel Vázquez Alonso
Facultad de Ciencias de la Educación, Universidad de las Islas Baleares

Introducción

Entre las principales conclusiones emanadas de los debates del grupo de trabajo sobre Ciencia, Tecnología y Sociedad (CTS) del 4º Simposio Internacional sobre Tendencias Mundiales de la Educación en Ciencia y Tecnología organizado por la IOSTE (*International Organization for Science and Technology Education*), celebrado en agosto de 1987 en el IPN de Kiel (Alemania), se manifestó una preocupación especial por la evaluación en este área de conocimiento, tanto del progreso en el aprendizaje de los estudiantes como de la enseñanza practicada por el profesorado y de los propios proyectos y programas CTS establecidos. Además, se estableció la urgente necesidad de investigar alternativas a la evaluación tradicional para suministrar nuevos criterios de evaluación, útiles para la enseñanza y el aprendizaje CTS, así como la búsqueda de instrumentos adecuados, válidos y fiables, para aprendizajes que no son los tradicionales y normas de evaluación específicas para el marco de enseñanza CTS: actitudes, valores, etc. (Hofstein, Aikenhead y Riquarts, 1988).

En el mismo Simposio se criticó la insuficiente investigación realizada hasta entonces en este aspecto de la educación CTS, lo que resulta particularmente grave, sobre todo porque los cursos CTS tratan de cuestiones polémicas y complejas debido a su propia naturaleza intrínsecamente dialéctica, siendo más difícil evaluar los logros de estos contenidos que los éxitos en los aprendizajes requeridos por los cursos más convencionales de ciencia. Así mismo, se insistió mucho en que la evaluación de los estudiantes tiene que ser coherente con los objetivos propuestos en tales cursos CTS. El grupo de trabajo CTS también reclamó la necesidad de comparar la calidad de los *items* de opción múltiple contruidos teóricamente con los desarrollados empíricamente (Aikenhead, 1988). Por último, se concluyó que, hasta el momento de la celebración del Simposio, se carecía de suficiente fundamentación teórica para establecer una estructura útil para la evaluación CTS, por lo que estas cuestiones deberían constituir un tema de interés prioritario para el futuro.

La evaluación de las actitudes puede suponer un importante obstáculo curricular a la hora de afrontar la innovación que supone la educación CTS, pues pocos profesores están dispuestos a incorporar temas nuevos en sus aulas sin tener una idea clara de cómo evaluarlos. Como muestra de ello puede señalarse que en un estudio longitudinal realizado para desarrollar y actuar sobre las ideas acerca de la naturaleza de la ciencia (una de las dimensiones CTS importantes) con profesores en formación inicial, se comprobó que éstos eran capaces de adquirir una adecuada comprensión sobre diversos aspectos de la

naturaleza de la ciencia y que muchos de ellos llegaban a enseñarlos explícitamente en clase, pero no se interesaron ni se esforzaron demasiado por evaluar lo que los estudiantes aprendían sobre estos temas, lo que de alguna manera refleja la dificultad de su evaluación (Bell, Lederman y Abd-El-Khalick, 2000).

Desde una posición crítica, este artículo se ocupa de cuestiones metodológicas ligadas a la investigación de las actitudes y creencias CTS del alumnado y el profesorado (en formación inicial o en ejercicio), informándose de los progresos más recientes en este campo, sobre todo de los que se refieren a la elaboración y desarrollo de nuevos instrumentos de evaluación y los procedimientos a seguir para la valoración de las respuestas: modelos de respuestas, métricas, etc.

Instrumentos para la evaluación de actitudes y creencias CTS

El análisis de la validez y fiabilidad de los numerosos instrumentos de evaluación y técnicas utilizados para realizar los estudios sobre las actitudes y creencias CTS de los estudiantes y del profesorado, conforma toda una línea de investigación en sí misma dentro de la didáctica de las ciencias experimentales. Al igual que ha ocurrido en otras ramas de la investigación educativa, con el paso de los años la investigación sobre las actitudes y creencias CTS se ha ido desplazando lentamente desde los enfoques de evaluación cuantitativos a otros más cualitativos. Los primeros trabajos solían utilizar escalas de Likert, diferenciales semánticos y cuestionarios o tests de opción múltiple, generalmente diseñados desde la óptica del investigador. Estudios posteriores con entrevistas clínicas revelaron discrepancias entre lo que entendían los estudiantes y lo que los investigadores decían en los enunciados de los cuestionarios, circunstancia a la que Aikenhead se refiere irónicamente como la "inmaculada percepción". En consecuencia, se pusieron en cuestión muchas de las conclusiones de las primeras investigaciones, porque casi ninguna de ellas utilizó entrevistas clínicas semi-estructuradas para corroborar los datos obtenidos con instrumentos de papel y lápiz. El siguiente es un ejemplo bastante drástico de todo esto:

En un estudio llevado a cabo por Lederman y O'Malley (1990) con una muestra de 69 estudiantes de educación secundaria y bachillerato, se administró un cuestionario con cuatro preguntas abiertas para evaluar algunas creencias sobre la naturaleza de la ciencia. El cuestionario se pasó al comienzo (pre-test) y al final del año escolar (pos-test). Después de clasificar las respuestas según mostraran un punto de vista "absoluto" o "provisional" del conocimiento científico, se seleccionó una muestra estratificada (teniendo en cuenta las creencias, los cambios observados entre el pre-test y el pos-test, y el nivel escolar de los alumnos) para hacer entrevistas que fueron grabadas en video. Durante el transcurso de éstas, se preguntó a los estudiantes sobre sus respuestas escritas con el fin de aclararlas. Mientras que las respuestas a los cuestionarios indicaban que la mayoría del alumnado tenía un punto de vista "absoluto" (esto es, pensaban que el conocimiento científico correcto básicamente no cambia), las entrevistas mostraron con bastante claridad que en realidad sus creencias tendían más bien hacia la provisionalidad del conocimiento científico. Los autores concluyeron su estudio manifestando que era esencial el uso de entrevistas clínicas para evaluar la comprensión de los estudiantes sobre la naturaleza de la ciencia. Así mismo, señalaron muchas de las deficiencias que aparecen cuando los investigadores intentan clasificar las respuestas escritas de los estudiantes a los clásicos tests de papel y lápiz.

Aunque la mayoría de las investigaciones cualitativas realizadas con entrevistas clínicas no han contradicho básicamente los resultados obtenidos anteriormente con enfoques cuantitativos que usaban cuestionarios, sí que han permitido hacer evaluaciones más finas y válidas de las creencias de los estudiantes y los profesores (Acevedo, 2000a), proporcionando a los investigadores un conocimiento más matizado del tema y una visión más contextualizada de las secuencias de enseñanza y los factores que median en las propias concepciones (Lederman, 1992).

Es de sobra conocido que los problemas metodológicos constituyen una fuente permanente de debates entre los investigadores en educación, en general, y de didáctica de las ciencias, en particular. Hace poco más de una década, se destacaba que las técnicas cualitativas en la investigación, sobre todo las entrevistas clínicas y los estudios de caso, permitirían a los investigadores evitar las dificultades que existen para encajar las respuestas escritas en categorías o puntos de vista establecidos *a priori*. El trabajo clásico de Aikenhead (1988) confrontando cuatro métodos para evaluar actitudes y creencias CTS (escalas de Likert, cuestionarios de opción múltiple, preguntas de respuesta abierta con párrafos breves y entrevistas clínicas semi-estructuradas), puso de manifiesto los problemas de las escalas de Likert y los cuestionarios clásicos, así como la mayor fiabilidad de las entrevistas. Aunque al compararlas con otros instrumentos de papel y lápiz las entrevistas clínicas suelen dar una información bastante segura, como contrapartida su utilización es muy costosa en tiempo de aplicación y esfuerzo de análisis, por lo que su empleo suele quedar necesariamente restringido a un número pequeño de personas (Acevedo, 2000a). Así pues, frente al aparente auge de las metodologías cualitativas, conviene no olvidar que tanto éstas como las cuantitativas tienen puntos débiles y fuertes, ventajas e inconvenientes, los cuales han sido señalados por especialistas en el tema, denunciando la "*kuhnificada*" polémica que simplifica en exceso el enfrentamiento maniqueo entre ambas metodologías, el cual acaba siempre señalando la falsa superioridad de una sobre otra (Manassero y Vázquez, 2000). La realidad es mucho más compleja y diversa, y exige "*deskuhnificar*" el debate (Shadish, 1995), rompiendo así con los estereotipos que se han generado y abriendo vías para la integración de ambas tradiciones metodológicas.

Afortunadamente, en la actualidad parece existir un acuerdo cada vez mayor en combinar las investigaciones cualitativas con las cuantitativas. No es de extrañar, por tanto, que en los últimos años hayan continuado los intentos de elaborar cuestionarios aplicables a muestras grandes, pero contruidos ahora empíricamente siguiendo las pautas del paradigma de la investigación cualitativa para no volver a caer en errores anteriores (en particular, el peligro potencial de que los estudiantes no perciban el mismo significado de los enunciados de los *items* que el investigador), de tal manera que sean instrumentos más válidos y fiables. En efecto, las opciones múltiples elaboradas empíricamente pretenden averiguar los puntos de vista de los estudiantes o los profesores que responden al cuestionario tratando de no influir con la perspectiva del investigador. Su principal característica es que las opciones múltiples recogen las opiniones expresadas en entrevistas piloto anteriores o las repuestas a cuestiones abiertas, escritas por personas semejantes a las que está destinado el cuestionario (Aikenhead y Ryan, 1989). Así pues, estos nuevos instrumentos constituyen una vía intermedia, que intenta aunar las ventajas de los cuestionarios (p.ej., su agilidad en la aplicación) con las de las entrevistas (p.ej., su fiabilidad y precisión).

El instrumento de estas características más completo para la investigación de concepciones, creencias y opiniones sobre las relaciones entre ciencia, tecnología y sociedad es posiblemente el inventario *Views on Science-Technology-Society* (VOSTS), preparado por Aikenhead, Fleming y Ryan (1987) y modificado posteriormente (Aikenhead y Ryan, 1992). Estos autores explican los pasos que siguieron para su desarrollo y discuten sobre su validez y fiabilidad (Aikenhead y Ryan, 1989). El VOSTS abarca temas correspondientes a las siguientes dimensiones conceptuales: definiciones de ciencia y tecnología, interacciones entre ciencia, tecnología y sociedad, sociología externa de la ciencia, sociología interna de la ciencia (características de científicos, construcción social de la tecnología y del conocimiento científico, ciencia y género) y naturaleza del conocimiento científico. Una posible clasificación de los tópicos CTS que se incluyen en las 114 cuestiones del VOSTS se muestra en el cuadro adjunto.

Otro cuestionario de opciones múltiples desarrollado empíricamente es el *Teacher's Belief about Science-Technology-Society* (TBA-STs), de Rubba y Harkness (1993). Se trata de un instrumento, mucho más breve, que sigue las pautas de construcción del VOSTS, pero elaborado para investigar las creencias del profesorado de ciencias.

Partiendo de la taxonomía de actitudes relacionadas con la ciencia y la tecnología propuesta por Vázquez y Manassero (1995) y tomando como referencia los instrumentos anteriores se ha construido el *Cuestionario de Opiniones sobre Ciencia, Tecnología y Sociedad* (COCTS), con un centenar de cuestiones adaptadas al contexto cultural español, que está destinado a evaluar las actitudes y creencias CTS; cuestionario disponible tanto en versión castellana (Vázquez y Manassero, 1997) como catalana (Manassero y Vázquez, 1998; Manassero, Vázquez y Acevedo, 2001).

Cuadro

Clasificación de tópicos CTS (Aikenhead y Ryan, 1992)

Definiciones
<p>1. <i>Ciencia y tecnología</i></p> <p>Concepto de ciencia (p.ej., instrumentalismo, satisfacer la curiosidad sobre el mundo, una empresa social).</p> <p>Concepto de tecnología (p.ej., intenciones humanas y sociales, aparatos, herramientas y máquinas, componentes socioeconómicos y culturales).</p> <p>Concepto de investigación y desarrollo (I+D).</p> <p>Interdependencia entre la ciencia y la tecnología (p.ej., rechazar la visión de la tecnología como ciencia aplicada).</p>

Epistemología

2. Naturaleza del conocimiento científico

Características de las observaciones científicas (p.ej., su carga teórica, los límites de la percepción).

Naturaleza de los modelos científicos y de los sistemas de clasificación.

Carácter provisional del conocimiento científico.

Hipótesis, teorías y leyes (p.ej., definiciones, papel de los supuestos, criterios para las opiniones).

Enfoques científicos de las investigaciones (p.ej., carácter no lineal, rechazo del procedimiento paso a paso, el "método científico" como una forma escrita).

Precisión e incertidumbre en el conocimiento científico y tecnológico (p.ej., razonamiento probabilístico).

Supuestos fundamentales para el conjunto de la ciencia (p.ej., universalidad).

Status epistemológico del conocimiento científico (p.ej., ontología como suposición, cuestionamiento del positivismo lógico).

Paradigmas *versus* universalidad de los conceptos en las diferentes disciplinas.

Sociología de la ciencia. Aspectos internos

3. Características de los científicos

Motivaciones personales de los científicos.

Valores y normas que orientan a los científicos en su trabajo y en su vida cotidiana (p.ej., mentalidad abierta, lógica, honestidad, objetividad, escepticismo, suspensión de opiniones; y los valores opuestos: mentalidad cerrada, subjetividad, etc.)

Ideología de los científicos (p.ej., punto de vista religioso).

Aptitudes necesarias para la ciencia (p.ej., compromiso, paciencia).

Efecto del género sobre los procesos y productos de la ciencia.

Representaciones subyacentes de las científicas.

4. Construcción social del conocimiento científico

Colectivización de la ciencia (p.ej., lealtades al equipo de investigación, a quien da el empleo).

Decisiones científicas (p.ej., desacuerdos entre los científicos, formas de conseguir el consenso).

Comunicación profesional entre científicos (p.ej., revisiones a fondo, revistas, publicaciones, conferencias).

Interacciones profesionales por la competitividad (p.ej., políticos, secretos, plagios).

Interacciones sociales.

Influencia de los individuos en el conocimiento científico.

Influencia nacional sobre el conocimiento científico y tecnológico.

Ciencia privada *versus* ciencia pública.

5. Construcción social de la tecnología

Decisiones tecnológicas.

Autonomía de la tecnología (p.ej., imperativo tecnológico).

Sociología de la ciencia. Aspectos externos

6. Influencia de la sociedad en la ciencia y la tecnología

Gobierno (p.ej., control sobre los recursos, política y actividades de la ciencia, influencia de los políticos).

Industria (p.ej., control corporativo dictado por los beneficios económicos).

Ejército (p.ej., utilización de los científicos como recursos humanos).

Ética (p.ej., influencia sobre los programas de investigación).

Instituciones educativas (p.ej., educación científica obligatoria).

Grupos de presión (p.ej., asociaciones de consumidores, organizaciones no gubernamentales, ecologistas, feministas).

Influencia de los ciudadanos sobre los científicos (p.ej., educación, interacciones sociales).

7. Influencia de la ciencia y la tecnología sobre la sociedad

Responsabilidad social de los científicos y los tecnólogos (p.ej., comunicación con el público, preocupaciones y responsabilidades sobre riesgos).

Contribución a las decisiones sociales (p.ej., toma de decisiones tecnocráticas *versus* democráticas, decisiones morales y legales, testimonios de los expertos, cabildeo para lograr fondos).

Creación de problemas sociales (p.ej., tomar en consideración las consecuencias negativas y positivas, competir por los fondos).

Resolución de problemas sociales y prácticos (p.ej., apoyo tecnológico, tipos de problemas cotidianos).

Contribución al bienestar económico (p.ej., riqueza y trabajo).

Contribución al poder militar.

Contribución al pensamiento social (p.ej., léxico, metáforas).

8. Influencia de la ciencia escolar sobre la sociedad

Establecer puentes entre las dos culturas (humanista y científico-tecnológica).

Poderes sociales (p.ej., decisiones de los consumidores).

Caracterización de la ciencia escolar.

Progresos recientes en la evaluación de las actitudes y creencias CTS

Desde hace relativamente poco tiempo, los esfuerzos se están dirigiendo a desarrollar nuevos modelos para la valoración de las respuestas de opción múltiple que aparecen en los cuestionarios desarrollados empíricamente. El formato de respuesta propuesto para el VOSTS, y utilizado hasta ahora en las evaluaciones hechas con el COCTS (Vázquez y Manassero, 1997), consiste en seleccionar una sola opción entre todas las alternativas que proporciona cada cuestión, aquella que mejor se ajusta a la opinión de la persona que responde. Este *Modelo de Respuesta Única* (MRU) es muy limitado, pues solo permite realizar comparaciones centradas en cada cuestión particular y no faculta para hacer comparaciones test-retest o la verificación de hipótesis, procedimientos que son básicos en la estadística inferencial.

En el marco del MRU, Rubba, Schoneweg-Bradford y Harkness (1996) sugieren puntuar la respuesta única de cada cuestión del VOSTS (o del TBA-STC) a partir de las tres categorías (Adecuada/Plausible/Ingenua) en las que clasifican las frases de las opciones múltiples del cuestionario, categorías que se establecieron después de ser evaluadas de modo independiente por siete jueces (ampliados a nueve cuando no había acuerdo en la clasificación). Las puntuaciones asignadas a cada categoría son: 3 (Adecuada), 2 (Plausible) y 1 (Ingenua). En tal caso, las puntuaciones acumuladas en varias cuestiones pueden utilizarse en procedimientos de estadística inferencial. Esta técnica supone un claro avance, pero también tiene importantes inconvenientes de medida porque las puntuaciones fijadas no discriminan respuestas que pueden ser claramente diferentes; de otra forma, el método de medida propuesto no tiene la fidelidad necesaria para representar adecuadamente las actitudes evaluadas, pues, por ejemplo, la misma medida podría corresponder a respuestas muy diferentes.

En general, el MRU aplicado para valorar las cuestiones de los instrumentos desarrollados empíricamente no permite obtener puntuaciones globales sobre el total de las preguntas, a modo de valoración ponderada total de los diversos aspectos que se representan en ellas. Solamente es posible una aproximación independiente a cada una de las cuestiones en un análisis individualizado tanto a escala general como para las diferencias entre grupos. Todas estas restricciones metodológicas son innecesarias porque se pueden proponer otras alternativas para potenciar más las posibilidades de utilización de estos cuestionarios en la investigación evaluativa de las actitudes y creencias CTS. Ahondar más en este análisis de la evaluación permite valorar con mayor precisión las actitudes y creencias sobre cada tema. Todo esto pasa por resolver los siguientes problemas (Vázquez y Manassero, 1999a):

- Emplear un nuevo modo de responder al COCTS, que deberá ser un *Modelo de Respuesta Múltiple* (MRM); es decir, un modelo donde las personas valoren todas las opciones de respuesta que se presentan en cada cuestión del COCTS (Vázquez, Manassero y Acevedo, 1998).
- Diseñar una métrica, es decir, un modelo de medida de las respuestas, que se plasme en un indicador cuantitativo (*índice actitudinal*) y adaptarlo a las necesidades del nuevo MRM adoptado para el COCTS (Vázquez y Manassero, 1999a,b).

- Aplicar baremos y clasificar detalladamente todas las opciones propuestas en las cuestiones del COCTS para poder aplicar la métrica diseñada (Manassero, Vázquez y Acevedo, 2001; Vázquez, Acevedo y Manassero, 2000).

Respecto del primer problema, hay que hacer notar que, además de las restricciones señaladas, el MRU tiene aún una limitación más importante, la de no utilizar gran parte de la información disponible en todas las opciones de cada cuestión; lo único que permite saber es que la alternativa elegida es la que más se ajusta a la opinión de la persona que responde, pero se desconoce su parecer sobre las demás. La información que se puede deducir con el MRU queda reducida, pues, a la opción elegida y su potencial adecuación respecto a los conocimientos actuales de epistemología, historia y sociología de la ciencia. Además, el MRU tiene el problema de la falta de control de las respuestas al azar, que es común a todos los cuestionarios de elección múltiple. La dificultad no reside tanto en la existencia de respuestas al azar como en su falta de control; es decir, no existe ninguna posibilidad de contrastar la autenticidad de la respuesta frente a otros elementos de información, porque éstos no existen. Para superar estos inconvenientes, manteniendo las ventajas métricas (fidelidad de las medidas y discriminación de respuestas) es evidente que un MRM resulta más adecuado, puesto que aporta muchos más datos, maximiza la información disponible en cada cuestión del COCTS y permite alcanzar el mayor grado de precisión en la medida del objeto actitudinal evaluado, por lo que se logrará tener una idea más exacta de las actitudes y creencias de cada persona.

En relación con el segundo problema, hay que hacer notar que el MRU está implícito en cualquier discusión que se haga sobre la métrica a utilizar. El origen de los defectos de la escala métrica de Rubba, Schoneweg-Bradford y Harkness (1996) está en la relación mutua entre las puntuaciones que proponen estos autores (Vázquez y Manassero, 1999b). Por razones de espacio, no es posible hacer aquí una discusión del problema¹. Baste con señalar que esta métrica puede superarse utilizando otra más discriminatoria para valorar cada una de las tres categorías de respuestas, asignando 3.5 (Adecuada), 1 (Plausible) y 0 (Ingenua).

El modelo empleado se enmarca en la tradición de la psicología social para la evaluación de actitudes y creencias conocida como *escalamiento psicofísico*, una alternativa a las técnicas de *evaluación psicométrica*. El escalamiento psicofísico se basa en graduar los estímulos aplicados a las personas evaluadas (p.ej., frases sobre creencias, afectos o conductas) en una dimensión psicológica, observando después sus respuestas o reacciones a esos estímulos para clasificarlas en un continuo de actitud. En cambio, la evaluación psicométrica se apoya en la aplicación de tests donde se responde a una serie de cuestiones, cada una de las cuales pretende valorar el atributo común que se trata de medir, a partir de cuyas respuestas se construye una puntuación que clasifica la actitud de las personas sobre un continuo. Las conocidas escalas tipo Likert y de diferencial semántico caerían dentro de esta segunda modalidad, cuya validez radica en la capacidad de cada cuestión para representar adecuadamente el constructo

¹ Manassero, Vázquez y Acevedo (2001) explican con detalle las principales limitaciones de la métrica de Rubba, Schoneweg-Bradford y Harkness, dando una alternativa razonada que permite superarlas (véanse también Vázquez y Manassero, 1999a,b; sobre todo el segundo artículo).

actitudinal que se pretende medir; validez que se ha dado siempre por supuesta ya que no ha habido preocupación entre los investigadores por confirmarla, de donde se han derivado la mayoría de los problemas métricos y defectos puestos de manifiesto anteriormente.

Para aplicar el MRM e interpretar las respuestas directas como medidas actitudinales se requiere el escalamiento o “baremación” previa de las opciones de cada cuestión, lo que nos lleva al tercero de los problemas señalados. Aunque no es posible entrar aquí en detalle sobre los entresijos metodológicos², cabe resaltar que la clasificación de las 637 opciones del COCTS en alguna de las tres categorías establecidas (Adecuada/Plausible/Ingenua) se ha realizado mediante la técnica de evaluación por un panel de jueces expertos. El procedimiento empleado es similar al de los modelos de escalamiento psicofísico de Thurstone o Guttman (Eagly y Chaiken, 1993), aunque no tan sofisticado estadísticamente porque la naturaleza del problema es muy diferente. En efecto, en el caso de las actitudes CTS los posibles objetos de actitud son muchos y variados (definición de ciencia, de tecnología, influencia social, características de los científicos, naturaleza del conocimiento científico, etc.), y, además, en algunos casos el acuerdo puede ser escaso y controvertido (Alters, 1997a,b; Eflin, Glennan y Reisch, 1999; Smith *et al.*, 1997; Vázquez, Acevedo, Manassero y Acevedo, 2001). Esta complejidad limita bastante la capacidad del escalamiento mediante jueces porque, en primer lugar, las actitudes relacionadas con la ciencia no son simples saberes comunes, sino que requieren conocimientos actualizados sobre diversas disciplinas (filosofía, sociología e historia de la ciencia, entre otras), de modo que no es fácil encontrar personas con la competencia necesaria para juzgar con validez y fiabilidad las opciones propuestas. Además, en segundo lugar, el gran número de cuestiones y el escaso margen para disponer de enunciados relevantes en las opciones de cada pregunta concreta, que están bastante predeterminadas, restringen tanto la posible selección estadística que resulta inoperante en este caso. Por último, la naturaleza controvertida y dialéctica de muchas cuestiones que tienen que ver con las actitudes y creencias CTS hace de la provisionalidad un rasgo inherente al escalamiento, lo que no debe entenderse, sin embargo, como ineficacia sino solo con cierto distanciamiento de cualquier sentido absoluto o definitivo sobre el tema, pendiente de nuevos juicios que mejoren la clasificación realizada (Vázquez, Acevedo y Manassero 2000).

Quizás podría reprocharse que el MRM propuesto reproduce una nueva versión de los clásicos cuestionarios tipo Likert. Estas escalas están formadas por un conjunto de frases cuyas puntuaciones se suman para obtener una puntuación total; su imprecisión radica precisamente en suponer que la escala refleja un constructo único, al cual supuestamente se deben ajustar todas las frases, hipótesis que casi nunca se cumplen. Muchas veces no se satisface porque el constructo no se ha definido, otras porque el que se ha definido no es unitario, sino multidimensional, esto es, está formado por diversos componentes muy diferentes; en consecuencia, no se puede saber qué representa exactamente la puntuación total (Gardner, 1996). Estas objeciones no son aplicables al MRM, porque el constructo está bien definido y las

² La descripción más completa del procedimiento metodológico seguido para clasificar las 637 opciones del COCTS se encuentra en Manassero, Vázquez y Acevedo (2001). También puede consultarse un resumen en Vázquez, Manassero y Acevedo (2000). Una comparación de los resultados del escalamiento de algunas cuestiones del COCTS con los obtenidos por Rubba y Harkness (1993) aparece en Vázquez, Acevedo y Manassero (2001).

puntuaciones se limitan a cada cuestión, donde la actitud se mide independiente e individualmente. Además, este procedimiento garantiza para cada cuestión la unidimensionalidad del constructo valorado, pues cada una de ellas se refieren al mismo tema de contenido. Por otro lado, las valoraciones directas en las escalas Likert originan de inmediato la puntuación correspondiente. En cambio, en el caso del MRM la puntuación directa se diversifica según la categoría asignada a esa posición a través del escalamiento, de modo que su contribución a la actitud global es más precisa por este método y su adecuación no es tan burda como la que resulta de tratar por igual todas las frases de las escalas Likert.

Con independencia de estas u otras objeciones metodológicas que se puedan aducir contra el MRM, no puede olvidarse tampoco que el MRU se obtiene de éste conservando la posición que obtiene la valoración de ajuste más alta y, en este único caso, el MRM añade aún una información adicional respecto al MRU, como es la de conocer la valoración asignada a la posición. Puede comprenderse fácilmente que cuando se añade información de todas las opciones, la actitud global así construida está más fundamentada, más informada y es más precisa.

No puede obviarse, sin embargo, que los procedimientos de valoración mediante jueces siempre tienen un margen de error; en primer lugar, debido a la variabilidad causada por el propio carácter dialéctico de los temas abordados en el COCTS, que convierte a algunos de ellos en polémicos, opinables, abiertos, etc., y, en segundo lugar, por el propio proceso de clasificación seguido, el cual metodológicamente implica el uso de parámetros estadísticos que, por sí mismos, añaden una nueva fuente de error. Todo esto hace que la clasificación obtenida sea necesariamente provisional y tentativa. Así mismo, el procedimiento de valoración por jueces realizado ha puesto de manifiesto la dificultad para obtener representadas las tres categorías establecidas en todas las cuestiones del COCTS; en particular, se han encontrado algunos casos que carecen totalmente de opciones adecuadas. Por tanto, la clasificación tendrá que perfeccionarse revisando los resultados del escalamiento y modificando la redacción de algunas cuestiones. Esta mejora puede conseguirse a partir de una profunda discusión de los desacuerdos observados entre los miembros del panel de jueces expertos (dieciséis inicialmente, que se redujeron a once en la versión final); debate que debería realizarse incluso a la luz de nuevos avances en historia, sociología y epistemología de la ciencia (Vázquez, Acevedo y Manassero, 2001). Éste podría ser el camino para lograr mayor fiabilidad en los resultados e incrementar la calidad global del COCTS como instrumento de evaluación y diagnóstico CTS.

Conclusiones y perspectivas futuras

Aunque de la lectura de este artículo pareciera desprenderse lo contrario, más de una década después del Simposio de Kiel, la evaluación de las actitudes y creencias CTS de los estudiantes (y el profesorado) se encuentra aún relativamente en precario y queda bastante por hacer, pese a la cada vez mayor importancia de la educación CTS en los *curricula* escolares de todo el mundo. Desde luego, esta deficiencia no es exclusiva de este campo de la didáctica de las ciencias experimentales, sino que se extiende a toda la investigación sobre evaluación. En efecto, en el *International Handbook of Science Education* (Fraser y Tobin, 1998), la sección dedicada a la evaluación en didáctica de las ciencias es algo menos amplia que las demás y en la primera línea se afirma que diez años antes no se habría podido

escribir. Además, la complejidad es mayor en un terreno donde las respuestas son todavía más abiertas, como corresponde a problemas en los que intervienen, junto a los aspectos más propiamente científicos y tecnológicos, otros de carácter social, cultural, moral, epistemológico, etc. La naturaleza intrínsecamente dialéctica de los temas CTS tiene como consecuencia la relativización, difuminación e incluso la pérdida de las referencias habituales y la necesidad de crear marcos alternativos que suministren nuevos criterios para evaluar las cuestiones CTS.

Un ejemplo de otro aspecto poco tratado en las investigaciones sobre la evaluación de creencias y actitudes CTS es el siguiente. A pesar de la reciente concepción del mundo como *aldea global*, las respuestas a cuestiones CTS del tipo de las que se proponen en el VOSTS y el COCTS están a menudo ligadas al contexto, porque las percepciones de las personas sobre estos temas suelen depender en buena medida de normas sociales, culturales y políticas que en muchos casos tienen un marcado carácter local, regional o nacional (Acevedo, 1995, 1996b). En la bibliografía sobre el tema pueden encontrarse pruebas de diferencias interculturales al comparar los resultados de cuestiones comunes obtenidos en distintos países; por ejemplo, los de Zoller *et al.* (1990, 1991a,b) logrados en el sistema educativo canadiense (angloparlante) con los de Ben-Chaim y Zoller (1991; Zoller y Ben-Chaim 1994) conseguidos en el sistema educativo israelita, ya que en ambos casos se utilizan las mismas cuestiones del VOSTS e idéntica metodología de investigación. En consecuencia, muchas de las creencias y actitudes CTS detectadas, solamente pueden tomarse como tendencias de opinión que habría que contrastar investigando en diversos países del mundo, con especial incidencia en los iberoamericanos, para tratar de catalogar semejanzas y diferencias en las creencias y actitudes CTS. El siguiente paso para avanzar en la solución de este problema, podría ser la realización de meta-análisis de los trabajos que se han hecho últimamente (y se hagan en el futuro) en los diversos sistemas educativos. Por otro lado, también hay que tener en cuenta este carácter local de las actitudes y creencias CTS a la hora de desarrollar empíricamente muchas de las cuestiones que se utilizan en los instrumentos de evaluación, lo que no es ni mucho menos una cuestión trivial.

Por último, es preciso hacer hincapié en la necesidad de continuar progresando en el desarrollo de instrumentos adecuados para evaluar las creencias y actitudes CTS. La utilización de cuestionarios elaborados empíricamente (como, p.ej., los aquí señalados: VOSTS, TBA-STs y COCTS), que permitan al mismo tiempo valorar las diversas cuestiones de acuerdo con un modelo de respuesta múltiple y una métrica adecuada, serviría para aunar las ventajas de las evaluaciones que siguen las tradiciones de investigación cualitativa y cuantitativa, así como para atemperar las desventajas propias de cada una de ellas. Evaluar creencias y actitudes CTS con instrumentos y criterios de esta clase, que tratan de reducir la influencia de las percepciones del investigador y dan la máxima información posible sobre cada pregunta, permitirá, sin duda, seguir avanzando en este importante campo de investigación de la didáctica de las ciencias experimentales.

Referencias bibliográficas

- ACEVEDO, J.A. (1995). Educación tecnológica desde una perspectiva CTS. Una breve revisión del tema. *Alambique*, 3, 75-84.
- ACEVEDO, J.A. (1996b). La tecnología en las relaciones CTS. Una aproximación al tema. *Enseñanza de las Ciencias*, 14 (1), 35-44.
- ACEVEDO, J.A. (2000a). Algunas creencias sobre el conocimiento científico de los profesores de Educación Secundaria en formación inicial. *Bordón*, 52(1), 5-16.
- ACEVEDO, J.A. (2000b). *Evaluación de creencias sobre Ciencia, Tecnología y Sociedad en Educación*. Conferencia impartida en las I Jornadas Universitarias de Nerva: *Ciencia, Tecnología y Humanismo en la Sociedad Actual*. Concejalía de Educación del Excelentísimo Ayuntamiento de Nerva y Universidad de Huelva.
- AIKENHEAD, G.S. (1988). An analysis of four ways of assessing student beliefs about STS topics. *Journal of Research in Science Teaching*, 25(8), 607-629.
- AIKENHEAD, G.S., FLEMING, R.W. y RYAN, A.G. (1987). High-school graduates' beliefs about science-technology-society. I. Methods and issues in monitoring student views. *Science Education*, 71(2), 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. Saskatoon (Canadá): Department of Curriculum Studies, University of Saskatchewan.
- AIKENHEAD, G.S. y RYAN, A.G. (1992). The development of a new instrument: 'Views on Science-Technology-Society' (VOSTS). *Science Education*, 76(5), 477-491.
- ALTERS, B.J. (1997a). Whose nature of science? *Journal of Research in Science Teaching*, 34(1), 39-55.
- ALTERS, B.J. (1997b). Nature of science: a diversity or uniformity of ideas? *Journal of Research in Science Teaching*, 34(10), 1105-1108.
- BELL, R. L., LEDERMAN, N. G., ABD-EL-KHALICK F. (2000). Developing and acting upon one's conception of the nature of science: A follow-up study. *Journal of Research in Science Teaching*, 37(6), 563-581.
- BEN-CHAIM, D. y ZOLLER, U. (1991). The STS outlook profiles of Israeli High School students and their teachers. *International Journal of Science Education*, 13(4), 447-458.
- EAGLY, A.H. y CHAIKEN, S. (1993). *The psychology of attitudes*. Forth Worth: Harcourt Brace College Publishers.
- EFLIN, J.T., GLENNAN, S. y REISCH, R. (1999). The Nature of Science: A Perspective from the Philosophy of Science. *Journal of Research in Science Teaching*, 36(1), 107-116.
- FRASER, B.J. y TOBIN, K.G., Eds. (1998). *International Handbook of Science Education*. Dordrecht: Kluwer Academic Publishers.
- GARDNER, P.L. (1996). The dimensionality of attitude scales: a widely misunderstood idea. *International Journal of Science Education*, 18, 913-919.
- HOFSTEIN, A., AIKENHEAD, G. y RIQUEARTS, K. (1988). Discussions over STS at the Fourth IOSTE Symposium. *International Journal of Science Education*, 10(4), 357-366.
- LEDERMAN, N.G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, 29(4), 331-359.

- LEDERMAN, N.G. y O'MALLEY, M. (1990). Students' perceptions of tentativeness in science: Development, use, and sources of change. *Science Education*, 74(2), 225-239.
- MANASSERO, M.A. y VÁZQUEZ, A. (1998). *Opinions sobre ciència, tecnologia i societat*. Palma de Mallorca: Conselleria d'Educació, Cultura i Esports.
- MANASSERO, M.A. y VÁZQUEZ, A. (2000). Creencias del profesorado sobre la naturaleza de la ciencia. *Revista Interuniversitaria de Formación del Profesorado*, 37, 187-208.
- MANASSERO, M.A., VÁZQUEZ, A. y ACEVEDO, J.A. (2001). *Avaluació de temes de ciència, tecnologia i societat*. Palma de Mallorca: Conselleria d'Educació i Cultura.
- 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(4), 407-431.
- RUBBA, P.A., SCHONEWEG-BRADFORD, C.S. y HARKNESS, W.J. (1996). A new scoring procedure for the Views on Science-Technology-Society instrument. *International Journal of Science Education*, 18(4), 387-400.
- SHADISH, W.R. (1995). The quantitative-qualitative debates: 'Deskunifying' the conceptual context. *Evaluation and Program Planning*, 18, 47-49.
- 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), 1101-1103.
- 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 I.P. Martins (Coord.): *O Movimento CTS na Península Ibérica. Seminário Ibérico sobre Ci_ncia- Tecnologia-Sociedade no ensino-aprendizagem das ci_ncias experimentais*, pp. 219-230. Aveiro (Portugal): Universidade de Aveiro.
- VÁZQUEZ, A., ACEVEDO, J.A. y MANASSERO, M.A. (2001, en prensa). Enseñando ciencia: consenso y disenso en la educación y la evaluación de actitudes relacionadas con la ciencia. *XIX Encuentros de Didáctica de las Ciencias Experimentales*. Madrid: Universidad Complutense.
- VÁZQUEZ, A., ACEVEDO, J.A., MANASSERO, M.A. y ACEVEDO, P. (2001). Cuatro paradigmas básicos sobre la naturaleza de la ciencia. *Argumentos de Razón Técnica* (aceptada su publicación).
- VÁZQUEZ, A. y MANASSERO, M.A. (1995). Actitudes relacionadas con la ciencia: una revisión conceptual. *Enseñanza de las Ciencias*, 13(3), 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. (1999a). Evaluación educativa de los temas Ciencia-Tecnología-Sociedad. En C. Martínez Losada y S. García Barros (Eds): *La Didáctica de las Ciencias. Tendencias actuales*, pp. 675-685. La Coruña: Servicio de Publicaciones de la Universidad de La Coruña.
- VÁZQUEZ, A. y MANASSERO, M.A. (1999b). Response and scoring models for the 'Views on Science-Technology-Society' Instrument. *International Journal of Science Education*, 21(3), 231-247.
- VÁZQUEZ, A., MANASSERO, M.A. y ACEVEDO, J.A. (1998). Modelos y cuestiones de evaluación en Ciencia, Tecnología y Sociedad (CTS). *Resúmenes de los XVIII Encuentros de Didáctica de las Ciencias Experimentales*, pp. 239-240. La Coruña: Universidad de La Coruña.
- ZOLLER, U. y BEN-CHAIM, D. (1994). Views of Prospective Teachers Versus Practising Teachers about Science, Technology and Society Issues. *Research in Science & Technological Education*, 12(1), 77-89.

ZOLLER, U., DONN, S., WILD, R. y BECKETT, P. (1991a). Students' versus their teachers' beliefs and positions on science/technology/society-oriented issues. *International Journal of Science Education*, 13(1), 25-36.

ZOLLER, U., DONN, S., WILD, R. y BECKETT, P. (1991b). Teachers' beliefs and views on selected science-technology-society topics: A probe into STS literacy versus indoctrination. *Science Education*, 75(5), 541-561.

ZOLLER, U., EBENEZER, J., MORELY, K., PARAS, S., SANDBERG, V., WEST, C., WOLTERS, T. y TAN, S.H. (1990). Goal attainment in science-technology-society (S/T/S) education and reality: The case of British Columbia. *Science Education*, 74(1), 19-36.