

Caballo

6. LOS CUESTIONARIOS DE AUTOINFORME EN LA EVALUACIÓN CONDUCTUAL

BERNARD J. JENSEN

1. INTRODUCCIÓN

Este capítulo trata sobre los cuestionarios clínicos de autoinforme en la evaluación conductual. Su objetivo es proporcionar una revisión de una serie de cuestiones importantes para el uso apropiado de este enfoque de evaluación.

Numerosos autores han intentado resumir las diferencias entre los enfoques de evaluación conductual y tradicional (Goldfried y Kent, 1972; Mash y Terdal, 1976; Hartmann, Roper y Bradford, 1979; Nelson y Hayes, 1986; Barrios, 1988). Sin embargo, conforme la terapia de conducta ha crecido y cambiado en los últimos 30-35 años, los límites entre los enfoques conductual y no conductual se han vuelto cada vez más "permeables" (Haynes, 1984).

Como consecuencia, a veces es difícil discriminar entre cuestionarios "conductuales" y "no conductuales". Con seguridad, muchos de los cuestionarios desarrollados y utilizados por los terapeutas de orientación conductual son más específicos a la situación que los inventarios de personalidad tradicionales como el *Minnesota multiphasic personality inventory* (MMPI). Además, los formatos de opción de respuesta en muchos cuestionarios "conductuales" permiten que los pacientes seleccionen entre opciones conductuales específicas, de modo consistente con un constructo conductual.

Sin embargo, ¿concluiríamos que el *Beck depression inventory*, BDI [Inventario de depresión de Beck] (Beck y cols., 1961) es un cuestionario conductual? Por una parte, el BDI lo utilizan frecuentemente los terapeutas de conducta que trabajan con pacientes deprimidos y normalmente se incluye y se describe en los capítulos que tratan sobre la evaluación conductual de la depresión (Carson, 1986; Rehm, 1988). Pero, por otra parte, dicho cuestionario se desarrolló originalmente para ser administrado por entrevistadores clínicos, se centra únicamente en determinados síntomas (especialmente cognitivos) de la depresión y no es, en absoluto, situacionalmente específico.

De este modo, puede que muchos de los cuestionarios que se utilizan hoy día en la evaluación conductual no se ajusten siempre a los supuestos de un

University of Central Florida (EE UU).

Me gustaría expresar mi agradecimiento al profesor V. E. Caballo por su traducción del capítulo y a Jennifer Irvin por su ayuda en la preparación del manuscrito.

constructo conductual. Afortunadamente, cuando los evaluadores conductuales utilizan cuestionarios de autoinforme, dichos instrumentos no constituyen las únicas fuentes de información, ya que la evaluación a través de diversos métodos se ha convertido en la norma. Como consecuencia, la información obtenida a partir de los cuestionarios de autoinforme puede compararse con, o clarificarse por, la información resultante de otros enfoques de evaluación. Por consiguiente, aunque la metodología de los cuestionarios de autoinforme se ha sacado de enfoques de evaluación clínica más tradicionales y las distinciones entre distintas clases de cuestionarios (conductual frente a no conductual) son difusas, los cuestionarios de autoinforme pueden jugar un importante papel en las múltiples y variadas funciones de la evaluación conductual.

En el presente capítulo se expondrá en primer lugar una revisión histórica del empleo de los cuestionarios de autoinforme en la evaluación conductual. A continuación, se describirá toda una serie de ventajas asociadas a los cuestionarios de autoinforme. Después de justificar la relevancia de los enfoques de evaluación nomotética en la evaluación conductual, se revisarán distintos conceptos psicométricos tradicionales (por ejemplo, fiabilidad y validez). Finalmente, se realizará una discusión sobre cómo los cuestionarios de autoinforme pueden cumplir varias funciones en la evaluación conductual (por ejemplo, selección de conductas objetivo, diseños de programas de intervención y evaluación de la efectividad del tratamiento), incluyendo también varias sugerencias para mejorar el empleo de los cuestionarios.

II. REVISIÓN HISTÓRICA

Durante los 10 ó 15 primeros años de la historia de la terapia de conducta, los clínicos mostraron poco interés por desarrollar y emplear los cuestionarios como fuentes básicas para obtener información relativa a la evaluación. Los terapeutas de conducta preferían métodos de evaluación más "objetivos", confiando especialmente en la observación de la vida real como una fuente esencial para conseguir información relativa a la evaluación. Relacionado con lo anterior estaba también que los terapeutas de conducta mostraban una notable preferencia por los métodos de evaluación "directos" frente a los "indirectos". Los métodos que evalúan las conductas problema en el momento y en el lugar de su ocurrencia real se dice que son métodos directos, mientras que los métodos que lo hacen en otro tiempo y/o en otro lugar se denominan métodos indirectos (Barrios, 1988). Los cuestionarios (junto con las entrevistas y las evaluaciones retrospectivas por los demás) constituyen métodos indirectos de evaluación.

Finalmente, los terapeutas de conducta tienden a asumir que los cuestionarios de autoinforme (y otros métodos subjetivos de autoinforme como la entrevista) son susceptibles, de forma inherente, a un gran número de factores de

sesgo. Entre estos factores se encuentran el olvido, el engaño y la deseabilidad social (Anastasi, 1988).

Además de estos aspectos negativos, otra variable que contribuyó al uso limitado de los cuestionarios en la primera etapa de la evaluación conductual fueron las suposiciones de la teoría de los rasgos referentes al uso de los cuestionarios. A lo largo de la historia de la psicología, los teóricos de la personalidad han desarrollado numerosos cuestionarios para evaluar la "cantidad" que posee un individuo de un determinado rasgo (por ejemplo, el *locus* de control interno o la necesidad de logro). Sin embargo, históricamente, los terapeutas de conducta han sido bastante críticos con estas posiciones, prefiriendo adoptar puntos de vista situacionales (o al menos interaccionistas) sobre la conducta. De este modo, el énfasis de la evaluación conductual se ha puesto en la evaluación de lo que una persona hace en vez de evaluar lo que una persona tiene (Mischel, 1968).

Aunque pueden existir todavía estos sesgos contra los cuestionarios entre algunos terapeutas de conducta, no hay duda de que la práctica de la terapia de conducta en la clínica ha adoptado progresivamente el uso de una amplia variedad de medidas de autoinforme. La entrevista, los procedimientos de autorregistro y los cuestionarios de autoinforme se encuentran entre los enfoques de evaluación más frecuentemente utilizados por los terapeutas de conducta clínicos en Estados Unidos (Swan y MacDonald, 1978; Wade, Baker y Hartmann, 1979).

Está fuera del alcance de este capítulo enumerar la gran cantidad de cuestionarios que se han desarrollado para ayudar a los terapeutas de conducta contemporáneos a evaluar y tratar toda una serie de problemas. Sin embargo, otros capítulos de este libro servirán, sin duda, para apoyar la cuestión general de la proliferación relativa al desarrollo y empleo de cuestionarios. Se anima al lector interesado en el uso de estos instrumentos con problemas específicos (por ejemplo, depresión, ansiedad, trastornos alimentarios, etc.) a que consulte los capítulos específicos de este manual y de otras fuentes (Haynes y Wilson, 1979; Barlow, 1981; Ciminero, Calhoun y Adams, 1986; Bellack y Hersen, 1988; Mash y Terdal, 1988) si quiere obtener una mayor información.

III. VENTAJAS DE LOS CUESTIONARIOS DE AUTOINFORME

Existen numerosas ventajas asociadas con el uso de los cuestionarios de autoinforme, y es probable que muchos de estos factores hayan contribuido directamente al uso cada vez mayor de los cuestionarios en la evaluación conductual. En primer lugar, los cuestionarios tienen una elevada razón coste-eficacia; se puede obtener una gran cantidad de información con una inversión mínima del tiempo del paciente o del terapeuta. Segundo, los cuestionarios son fáciles de puntuar de manera objetiva, eliminando la necesidad de inferencias por parte de

la persona que los pasa. Tercero, los cuestionarios utilizados en la evaluación conductual poseen normalmente "validez aparente" (es decir, el cuestionario parece medir aquello para lo que fue desarrollado), proporcionando una contribución positiva a la confianza y a la relación entre el clínico y el paciente. Cuarto, como ya se ha mencionado sucintamente, los cuestionarios se han desarrollado para evaluar una amplia variedad de problemas objetivo y pueden utilizarse para un gran número de poblaciones de pacientes. Quinto, los cuestionarios son potencialmente apropiados para evaluar los tres componentes empleados en la evaluación del triple sistema de respuesta.

En la evaluación conductual se ha enfatizado esta evaluación del triple sistema de respuesta, en contraste con las denominadas "torpes teorías" de la personalidad (Lang, 1968). En la evaluación del triple sistema de respuesta, el clínico está interesado en determinar los aspectos motores, fisiológicos y cognitivos de la conducta del paciente. Por ejemplo, en el contexto de una fobia específica, un clínico estaría interesado en obtener información sobre los síntomas motores (por ejemplo, temblar, ralentizar el paso o la conducta de evitación), los síntomas fisiológicos (aumento de la tasa cardíaca o quedarse sin respiración), y los síntomas cognitivos (por ejemplo, habla negativa con uno mismo o imágenes relativas a consecuencias peligrosas). Especialmente a la luz del reciente aumento de las influencias cognitivas en la terapia de conducta, se emplean los métodos de evaluación por medio del autoinforme (incluyendo, pero no limitándose, a los cuestionarios) debido a que proporcionan los medios principales para obtener información relativa a estos determinantes privados/internos de los problemas objetivo. Así, debido a todas las razones que acabamos de mencionar, es probable que los cuestionarios continúen siendo una importante fuente de información en la terapia de conducta.

IV. ENFOQUES IDIOGRÁFICOS FRENTE A ENFOQUES NOMOTÉTICOS

Una de las supuestas características distintivas de la terapia de conducta es su énfasis en el caso individual, el denominado enfoque "idiográfico". Este enfoque se ha comparado con enfoques tradicionales de evaluación que se centran en las diferencias entre los miembros de un grupo, el denominado enfoque "nomotético". Según algunos autores (Cone, 1981; Nelson, 1983), el enfoque nomotético de evaluación tiene poca relevancia para la terapia de conducta. Sin embargo, especialmente cuando tratamos con cuestionarios de autoinforme clínicos, tenemos que familiarizarnos y utilizar muchos de los principios desarrollados dentro de la tradición nomotética.

Como se ha señalado anteriormente, la utilización de cuestionarios de autoinforme es consistente con un enfoque más tradicional de evaluación, o lo que Cone (1988) denomina enfoque nomotético-de rasgos. Para algunos terapeutas de conducta (radicales) este enfoque puede ser difícil de aceptar, pero tal

como consideran Barrios y Hartmann (1986), muchos de los conceptos, temas y metodologías del enfoque de la evaluación tradicional tienen contribuciones que hacer a la evaluación conductual. De este modo, el conocimiento de muchos de los principios nomotéticos para la construcción y el empleo de cuestionarios extraídos de los paradigmas de la evaluación tradicional, son perfectamente apropiados para los terapeutas de orientación conductual. De forma clara, una de las perspectivas conceptuales y nomotéticas relativas a la utilización de cuestionarios implica a la "tradición psicométrica".

V. CUESTIONES PSICOMÉTRICAS

La psicometría se refiere a la aplicación de procedimientos matemáticos y de medición a los fenómenos psicológicos, en general, y a los tests mentales, en particular. La familiaridad con muchos de los principios psicométricos implicados en la construcción y análisis de cuestionarios permitirá al clínico emplear apropiada y satisfactoriamente los métodos de evaluación por medio del autoinforme. Entender estas cuestiones psicométricas ayudará al clínico a seleccionar cuestionarios más válidos y fiables, a utilizar de forma más apropiada los datos obtenidos por medio de ellos y, por consiguiente, a realizar la evaluación conductual de forma que tenga como resultado un tratamiento más eficaz.

V.1. *Fiabilidad*

La fiabilidad es un término utilizado para referirse a la estabilidad o a la consistencia de los resultados de la evaluación. Normalmente nos interesan dos tipos de fiabilidad en los cuestionarios de autoinforme: la consistencia interna y la fiabilidad test-retest.

La fiabilidad de la consistencia interna se refiere al grado en que los ítems de un cuestionario están relacionados entre sí. La fórmula que se ha utilizado más a menudo para calcular esa consistencia ha sido la técnica denominada fiabilidad de las dos mitades. En este enfoque, los ítems del cuestionario se dividen en dos partes (por ejemplo, ítems pares contra ítems impares) y se correlacionan las puntuaciones de estas dos mitades (utilizando a menudo la corrección Spearman-Brown). Cronbach (1951) desarrolló una fórmula para calcular el "coeficiente alfa", que es la media de todos los posibles coeficientes de correlación de las dos mitades. Este número es conceptualmente equivalente a la cantidad de varianza atribuible a factores comunes entre los ítems de los tests o, en otras palabras, el coeficiente alfa es un índice de la homogeneidad de los ítems del cuestionario. Puesto que muchos de los cuestionarios utilizados en la evaluación conductual se emplean para evaluar un único constructo (por ejemplo, asertividad), es esencial un elevado grado de homogeneidad. Además, incluso cuando

un cuestionario evalúa varios constructos relacionados usando subescalas separadas, es necesaria la homogeneidad dentro de los ítems de la subescala.

La segunda clase de fiabilidad que debe interesar a los clínicos de orientación conductual que utilizan cuestionarios es la fiabilidad test-retest. Ésta implica determinar el grado de consistencia de los resultados obtenidos al pasar los cuestionarios en una ocasión y en otra posterior. Normalmente, el intervalo de tiempo entre las dos ocasiones en las que se pasan los cuestionarios debe ser relativamente breve, ya que cuanto más largo sea el intervalo menos estables serán los resultados. Aunque la especificidad situacional de la conducta es una suposición básica de la terapia de conducta, muchos patrones de comportamiento muestran estabilidad temporal (Bem y Allen, 1974; Epstein, 1979, 1980). Por consiguiente, seleccionar cuestionarios de autoinforme con una fiabilidad test-retest demostrada es una consideración importante en la evaluación conductual.

V.2. Validez

En este contexto, la validez se refiere al grado en que un cuestionario evalúa aquello para lo que fue construido. Aunque es importante considerar la fiabilidad de los cuestionarios de autoinforme, la fiabilidad es un criterio necesario, pero no suficiente, para la validez (es decir, un test puede ser fiable pero no válido). A menudo se ha hablado en psicometría de varias clases de validez; dos de los aspectos más importantes de la validez son la validez de contenido y la validez en relación con el criterio. La validez de contenido se refiere al grado en que los ítems de un cuestionario muestrean el constructo que se está evaluando. A diferencia de las consideraciones (y demás cuestiones psicométricas) señaladas anteriormente con respecto a la fiabilidad, la validez de contenido se determina normalmente de manera subjetiva o cualitativa, en vez de un modo objetivo o cuantitativo. Una de las formas más frecuentes para determinar la validez de contenido consiste en preguntar a expertos en el campo que evalúen el grado en que los ítems incluidos en un cuestionario son representativos y apropiados. Este método es superior a fiarse simplemente de las impresiones de quien desarrolló el cuestionario, que podría pasar por alto ítems relacionados con áreas importantes o incluir ítems que no son relevantes.

La validez de contenido puede aumentarse obteniendo información de múltiples fuentes cuando se generan inicialmente ítems para un cuestionario. La primera parte del enfoque analítico conductual de Goldfried y D'Zurilla (1969), la fase del análisis situacional, proporciona un ejemplo de esta metodología. Al utilizar un análisis situacional se obtienen ítems de áreas de contenido relevante a partir de un gran número de sujetos objetivo, de sus iguales y otras personas significativas, y de profesionales del campo. El muestreo de un amplio rango de individuos puede ayudar a aumentar la probabilidad de que la validez de contenido del cuestionario sea elevada.

La validez en relación con el criterio se refiere al grado en que una medición

de un constructo (un predictor) está asociada con algún otro indicador del mismo constructo (el criterio). Las evaluaciones de la validez en relación con el criterio son especialmente importantes para los cuestionarios de autoinforme, puesto que las marcas hechas con bolígrafo por nuestros pacientes pueden estar relacionadas con las conductas criterio del "mundo real" que intentamos evaluar.

Hay dos tipos básicos de validez en relación con el criterio. La validez predictiva, que se refiere al grado de asociación entre las puntuaciones de una medida (por ejemplo, un cuestionario) y algún criterio independiente evaluado en una ocasión posterior. Por ejemplo, la validez predictiva de un cuestionario construido para evaluar la habilidad heterosocial podría evaluarse correlacionando las puntuaciones en el cuestionario con medidas posteriores del éxito en las citas. La validez concurrente, que se estudia cuando determinamos el grado de asociación entre nuestra medida y algún criterio independiente en el mismo momento. Por ejemplo, la validez concurrente de un cuestionario para evaluar las dificultades de comunicación de la pareja podría determinarse correlacionando las puntuaciones en ese cuestionario con una medida de observación de la comunicación marital obtenida durante la misma sesión de evaluación.

Aunque inicialmente puede no ser algo obvio, a menudo surge un problema práctico al seleccionar un criterio apropiado para los estudios sobre la validez relacionada con el criterio. Al comienzo de la historia de la terapia de conducta se suponía que los procedimientos de evaluación por medio de la observación eran inherentemente válidos y que, por consiguiente, estas medidas eran las más apropiadas como criterio. Sin embargo, la investigación ha revelado una serie de amenazas a la validez de los procedimientos de evaluación por medio de la observación (reactividad, sesgos del observador, errores de muestreo, etc.), dando como resultado un punto de vista más prudente sobre la validez de la información obtenido a través de los métodos de observación. Como consecuencia, si se encuentra una baja correlación entre las puntuaciones en el cuestionario y los datos de la observación, no se puede seguir suponiendo que la medida del cuestionario no es válida.

Cuando las puntuaciones de otros cuestionarios se utilizan como variable criterio, surge un problema adicional en los estudios sobre la validez de criterio. Aunque ésta puede ser una estrategia apropiada, especialmente cuando los cuestionarios criterio han estado ellos mismos sujetos a evaluaciones psicométricas rigurosas, a menudo se obtienen correlaciones significativas debido, en parte, al solapamiento de la varianza del método. Esto se refiere al hecho de que las correlaciones entre medidas pueden estar engañosamente infladas simplemente porque las medidas son del mismo tipo (por ejemplo, cuestionarios). Debemos, entonces, tener cuidado en asegurarnos de que la "validez en relación con el criterio" de nuestros cuestionarios esté evaluada adecuadamente.

V.3. Análisis factorial

El análisis factorial es un procedimiento estadístico para agrupar una serie de variables (en este caso, ítems de un cuestionario) en número menor de ellas que compartan la mayor cantidad de varianza común (Cattell, 1978). Uno de los propósitos que el análisis factorial cumple en la construcción y evaluación de los cuestionarios consiste en crear subescalas a partir del número total de ítems del cuestionario global.

Como se ha señalado anteriormente, al utilizar la suma de las puntuaciones de los cuestionarios, los terapeutas de conducta están adoptando implícitamente un punto de vista de la conducta humana basado en los rasgos. Así, una puntuación de 92 en un cuestionario que mide asertividad (por ejemplo, la *College self-expression scale*, *CSES* [Escala de autoexpresión universitaria] de Galassi, DeLo, Galassi y Bastien, 1974) puede interpretarse que significa que el individuo posee "92 unidades de asertividad". Sin embargo, por medio del análisis factorial se ha determinado que la *CSES* contiene ítems de aserción tanto "positiva" como "negativa". Así, el evaluador obtendría información específica más consistente con una perspectiva conductual de la evaluación, al saber que 60 de los 92 puntos provienen de los ítems de "aserción negativa", mientras que sólo 32 provienen de los ítems de "aserción positiva" (Caballo, 1993).

A menudo, si se han realizado investigaciones sobre el análisis factorial (y los factores pueden relacionarse también con la "estructura lógica"), raramente se emplea la puntuación total. Un ejemplo aplicado al área de los conflictos maritales es el *Relationship beliefs inventory*, *RBI* [Inventario de creencias sobre la relación] (Epstein y Eidelson, 1981; Eidelson y Epstein, 1982). El *RBI* contiene cinco subescalas que se obtuvieron por medio del análisis factorial; el empleo más frecuente del cuestionario se centra en las puntuaciones de las subescalas. En este enfoque, se pueden identificar y modificar las "ideas disfuncionales sobre la relación" específicas, en vez de utilizar una posición global para cambiar el pensamiento "irracional" (véase en este mismo libro el capítulo de Arias y Byrne).

V.4. Datos relativos a las normas

Como se ha señalado anteriormente, los cuestionarios cumplen funciones nomotéticas e idiográficas dentro de la evaluación conductual. Una de las maneras en las que se evoca la tradición nomotética es por medio de la referencia a la posición relativa de un individuo respecto a algún grupo sobre el que se conozcan las normas. Los evaluadores conductuales no siempre han estado interesados en este aspecto de los resultados de los cuestionarios, pero atender a los hallazgos relativos a las normas es útil para varios propósitos (Hartmann y cols., 1979; Barrios y Hartmann, 1986).

Los hallazgos de los cuestionarios de autoinforme pueden sugerir que la conducta del paciente es claramente deficiente o excesiva si se compararan los resultados con las normas del grupo. Además, se pueden determinar objetivos para el tratamiento o expectativas razonables para el cambio a partir de la información relativa a las normas. Igualmente, se pueden comparar sujetos de diferentes estudios o muestras con respecto a sus puntuaciones en los cuestionarios que han hecho referencia a las normas. Finalmente, se puede utilizar la información relativa a las normas como un criterio para las evaluaciones de validación social al mostrar que los cambios que resultan de la terapia son clínicamente significativos.

De este modo, es importante que aquellos que desarrollan nuevos cuestionarios para su uso en terapia de conducta recojan información relativa a las medias y las desviaciones típicas para muestras específicas. De igual manera, quienes utilizan cuestionarios deberían llegar a conclusiones sobre el significado de los resultados comparándolos con normas conocidas. Sin embargo, es muy importante asegurarse que la información relativa a las normas se utiliza de modo apropiado para distintas muestras de sujetos. Como ha señalado Anastasi (1988), las normas son relativas y flexibles, en vez de permanentes o absolutas. Por lo tanto, deben evaluarse las características de la muestra sobre la que se poseen normas cuando se realicen comparaciones y a menudo será necesaria la consideración de las normas del subgrupo específico. Muchos de los cuestionarios de orientación conductual que se emplean en Estados Unidos se han creado en ambientes académicos con estudiantes universitarios blancos, de clase media, como sujetos. Como consecuencia, el empleo de estos cuestionarios con poblaciones clínicas puede conducir a importantes errores de interpretación.

También tienen una relevancia especial las cuestiones culturales y lingüísticas. Conforme la sociedad de Estados Unidos se vuelve más variada racial y culturalmente, tenemos que asegurarnos que se lleven a cabo comparaciones con grupos apropiados sobre los que se posean normas. Además, traducir únicamente los cuestionarios de un idioma a otro y utilizar los "datos relativos a las normas" de la muestra de estandarización original puede llevar a errores en la evaluación.

VI. FUENTES DE ERROR EN LOS CUESTIONARIOS

Numerosas variables pueden afectar a la validez de la información obtenida por los cuestionarios. Entre estas variables se encuentran la discapacidad social, las características de la población, las características de "exigencia" de la situación de evaluación, las influencias reactivas, los sesgos de respuesta y otros factores asociados con el contexto de la administración de los cuestionarios o con los cuestionarios mismos. El grado en que el clínico comprenda e influya sobre estas variables tendrá un impacto sobre la fiabilidad de la información obtenida.

por los cuestionarios. Por consiguiente, se describirán brevemente cada uno de los factores anteriores.

Cuando se desarrollan cuestionarios para evaluar información socialmente relevante (por ejemplo, ajuste marital, abuso de sustancias, etc.), el deseo del paciente de aparecer con un tono positivo puede producir una información menos precisa. Estas influencias de deseabilidad social dan normalmente como resultado una mayor información de conductas positivas y una menor información de conductas negativas. También, como se ha señalado anteriormente, la mayoría de los cuestionarios empleados en la evaluación conductual posee validez aparente. Por consiguiente, los individuos que deseen presentar, a sabiendas, una falsa imagen de sí mismos pueden producir fácilmente información que disminuya la probabilidad de percepciones negativas, comprometiendo de esta manera la precisión del proceso de evaluación.

Las características de la población pueden también tener su impacto sobre la validez de los informes del cuestionario. Consistente con esta posición, se piensa que los miembros de determinados grupos de diagnóstico proporcionarán normalmente información poco fiable. Por ejemplo, entre los criterios diagnósticos para el trastorno de la personalidad antisocial se encuentra un desprecio de la verdad. Por consiguiente, es de esperar una elevada frecuencia de autoinformes inadecuados provenientes de individuos con este diagnóstico.

Una tercera variable que puede influir sobre la validez de la información obtenida por los cuestionarios implica a las características de exigencia de la situación de evaluación. Según Orne (1969), este concepto se refiere a los estímulos que influyen sobre la percepción de una persona con respecto a la naturaleza de una situación. Así, en un contexto de evaluación clínica, las personas que desean recibir compensación por sus problemas psicológicos pueden estar motivadas para sobreinformar acerca de su psicopatología, mientras que los que son remitidos por la justicia pueden comunicar menos problemas.

Otra fuente potencial de error en los datos de los cuestionarios implica las influencias reactivas; éstas se refieren al procedimiento mismo de evaluación que modifica la conducta objetivo que se está evaluando. Las influencias reactivas se han estudiado principalmente en el contexto de otros métodos de evaluación conductual (por ejemplo, autorregistro) y se ha dedicado poca atención a las posibles influencias reactivas asociadas con el uso de cuestionarios de autoinforme. Sin embargo, es poco probable que cuando se administran cuestionarios de modo regular, la conducta que se está evaluando cambie simplemente por razones relacionadas con la evaluación continua.

Una fuente de error adicional en la evaluación por medio de cuestionarios se centra en el sesgo de respuesta, que se refiere a la posibilidad de una dependencia consecutiva de las respuestas de un individuo a los ítems del cuestionario. Esto significa que una respuesta a un ítem puede influir sobre la respuesta al siguiente ítem y reducir así la validez de la información obtenida por el cuestionario. Dos recomendaciones específicas para disminuir la posibilidad del sesgo de respuesta implican alternar o variar los ítems que se formulan positivamente con

los que se formulan negativamente, y alternar o variar el orden de los valores numéricos que corresponden a opciones de respuesta (por ejemplo, en una escala de 7 puntos tipo Likert, empezar a veces con 1= "nunca", y a veces con 7= "siempre").

Además de las fuentes de error señaladas anteriormente, también pueden influir sobre la validez de los cuestionarios otras variables. Aspectos tales como las instrucciones, el formato de respuesta, la redacción de los ítems, las variables relativas al lugar donde se pasa el cuestionario, el número de instrumentos administrados simultáneamente, la exposición a experiencias de sensibilización y las interacciones previas con el evaluador conductual, pueden todos ellos contribuir a una disminución de la validez de la información obtenida por los cuestionarios.

Por lo que se ha visto, existen obviamente muchas fuentes de error que pueden tener su impacto sobre la validez o la precisión de la información obtenida por los cuestionarios. Sin embargo, sería erróneo concluir que porque alguna información de los cuestionarios no es válida, no lo es tampoco ninguna información recogida por ellos. Se necesita más investigación para controlar o reducir las fuentes de error.

VII. FUNCIONES DE LOS CUESTIONARIOS DE AUTOINFORME

Los terapeutas de conducta han utilizado los cuestionarios para una serie de propósitos. Tres de las principales funciones que cumplen son: 1) la identificación y selección de las conductas objetivo; 2) el diseño de programas de intervención; y 3) la evaluación de los resultados del tratamiento. Los cuestionarios sirven mejor para algunas de estas funciones que para otras y en los siguientes apartados se detallará de qué manera se han utilizado los cuestionarios. Además, se incluirán sugerencias con el fin de aumentar la relevancia de los cuestionarios para las funciones anteriores.

VII.1. Identificación y selección de las conductas objetivo

Uno de los objetivos principales de los cuestionarios conductuales consiste en identificar las conductas objetivo que se han de modificar. Como se ha señalado anteriormente, estos objetivos pueden incluir conductas motoras, fisiológicas y cognitivas que son problemáticas para el paciente. Debido a que se han desarrollado cuestionarios en un amplio rango de áreas, pueden ser especialmente válidos para la identificación o selección iniciales de las conductas problema.

Sin embargo, debido a que los terapeutas de conducta han puesto un gran énfasis sobre la especificidad al definir las conductas objetivo, normalmente se busca información relativa a la frecuencia, la intensidad, la duración y la magni-

tud de las distintas conductas. Desafortunadamente, los cuestionarios no proporcionan normalmente información con este grado de especificidad.

Con demasiada frecuencia, toda la información obtenida por un cuestionario es una única puntuación total. Supuestamente esta puntuación es un índice cuantitativo del grado en que una determinada área problema (por ejemplo, depresión) es relevante para un paciente, o la probabilidad de que el paciente emita una conducta que ha elegido entre varias (por ejemplo, aserción apropiada). Sin embargo, estas puntuaciones no proporcionan suficiente información específica relativa a los parámetros o a los determinantes situacionales de las conductas problema del paciente.

Una mayor atención a las puntuaciones de las subescalas del cuestionario podría mejorar esta situación. Muchos cuestionarios evalúan realmente varias dimensiones de una conducta problema y la utilización de una única puntuación resumen en un instrumento compuesto por muchos factores es inapropiada. Las puntuaciones de las subescalas pueden ser obtenidas de forma lógica o, aún mejor, de forma empírica (por ejemplo, por medio del análisis factorial), con el fin de proporcionar una información cuantitativa más específica, relevante para un punto de vista conductual de distintos trastornos. Por ejemplo, un índice general de "asertividad" es menos útil que índices separados de «rechazo de peticiones poco razonables», «expresión de emociones positivas» y «desacuerdo con el punto de vista de otra persona» (Caballo, 1993).

La especificidad de la información obtenida por los cuestionarios es también una función del tipo de opción de respuesta utilizado. Los evaluadores conductuales han empleado cuestionarios con una variedad de formato de respuesta, incluyendo listas, opciones verdadero-falso, respuestas a lo largo de una escala y afirmaciones con las que se está o no de acuerdo. Con el fin de que los cuestionarios proporcionen información más específica relativa a la frecuencia, duración, gravedad, intensidad o magnitud de distintas conductas, las opciones de respuesta deberían reflejar este énfasis específico.

Sería también beneficioso un aumento de la atención sobre temas de validez de contenido, que tengan en cuenta la naturaleza multimodal de los problemas de conducta. Si utilizamos ejemplos del área de la depresión, cuestionarios que son muy empleados actualmente, como el *BDI* o la *Depression adjective checklist* [Lista de adjetivos sobre la depresión] de Lubin (1967), muestrean información relativa a solamente uno o dos aspectos de la conducta deprimida (por ejemplo, síntomas cognitivos o afectivos) (Rehm, 1981, 1988). Si aumentamos el número de áreas evaluadas por un cuestionario (por ejemplo, los patrones de interacción social, los refuerzos sociales de la conducta deprimida), se podría obtener un análisis más amplio y comprensivo de estas conductas.

A pesar de las limitaciones de los cuestionarios de autoinforme como herramientas para identificar y seleccionar las conductas meta resaltadas por estas sugerencias, los cuestionarios cumplen claramente ese objetivo como componentes de un enfoque de evaluación multimodal.

VII.2. Diseño de programas de intervención

Una segunda función de los cuestionarios de autoinforme en la evaluación conductual es ayudar al diseño de programas de intervención, existiendo una serie de factores relevantes para esta compleja tarea de toma de decisiones; entre éstos se encuentran los siguientes: las conductas objetivo seleccionadas, las preferencias subjetivas del terapeuta, la información que se obtiene del paciente y los hallazgos de la investigación relativos a la eficacia comparada de diferentes programas de intervención (Cormier y Cormier, 1991). Dos factores concretos pueden ser importantes para seleccionar enfoques específicos de tratamiento: la comprensión de las variables que controlan los problemas objetivo y la aceptabilidad de distintos enfoques de intervención. Los cuestionarios pueden jugar un papel en proporcionar información relativa a estos dos factores.

Los terapeutas de conducta han puesto históricamente un gran énfasis en el determinismo ambiental y en el control situacional de la conducta (Mischel, 1968; Goldfried y Linehan, 1977; Haynes, 1983). Sin embargo, el concepto de determinismo ambiental estricto ha sido reemplazado por un modelo interaccionista que enfatiza el determinismo recíproco (Bandura, 1977, 1978). Este enfoque reconoce que no sólo el ambiente influye sobre la conducta del individuo, sino que la conducta del individuo cambia también el ambiente. Aunque existen varios modelos, el modelo E-O-R-C de Goldfried y Sprafkin (1976) considera la interacción entre variables cognitivas, emocionales y fisiológicas complejas que controlan la conducta.

La información sobre estas variables controladoras se puede obtener por medio de cuestionarios. Sin embargo, el nivel actual de sofisticación para identificar la relación entre respuestas conductuales y variables controladoras por medio de los cuestionarios de autoinforme no es precisamente grande. Una de las limitaciones del empleo de cuestionarios para identificar las variables controladoras es su insuficiente especificidad. Se recomienda que los ítems de los cuestionarios incluyan más información sobre posibles factores situacionales, cognitivos, emocionales y fisiológicos que controlan los problemas de conducta. De esta forma, los programas de intervención pueden centrarse en la modificación de aquellas variables que sean más importantes para las conductas objetivo identificadas. Por ejemplo, en el grado en que un cuestionario revele la presencia de procesos cognitivos desadaptativos, se pueden seleccionar estrategias de intervención que modifiquen estos procesos y las conductas problema con las que covarían.

Una segunda variable que influye sobre el diseño de los programas de tratamiento es la aceptabilidad de potenciales enfoques de intervención. La aceptabilidad del tratamiento se refiere a las evaluaciones subjetivas hechas por los pacientes sobre variables tales como la eficacia y la adecuación de las estrategias de intervención. Kazdin (1977) y Wolf (1978) creen que no es suficiente que las intervenciones conductuales sean eficaces, también tienen que ser muy aceptables.

La aceptabilidad del tratamiento afecta supuestamente a la continuación de la terapia, a la adherencia a las estrategias de intervención y a la satisfacción global con el tratamiento. De este modo, uno de los factores que influyen sobre el diseño de los programas de intervención incluye los juicios sobre la aceptabilidad del tratamiento.

La literatura sobre las reacciones de los pacientes (o potenciales pacientes) ante distintos enfoques de intervención está aumentando, surgiendo cada vez más información en las áreas de adelgazamiento en niños (Frentz y Kelley, 1986), intervenciones con ancianos (Lundervold, Lewin y Bourland, 1990) y terapia de pareja conductual (Upton y Jensen, 1991). Virtualmente, en todas estas intervenciones, los cuestionarios de autoinforme sirven como los principales métodos para evaluar la aceptabilidad del tratamiento (véase Elliot, 1988, para una revisión). Sin embargo, los cuestionarios específicos empleados con estos fines no son, a menudo, psicométricamente adecuados, debiéndose mejorar el diseño y el uso de los cuestionarios de autoinforme para la evaluación de la aceptabilidad del tratamiento.

VII.3. *Evaluación del tratamiento*

Finalmente, otra de las funciones principales de la evaluación conductual es evaluar el impacto de los programas de intervención sobre las conductas objetivo. Tanto si se emplean diseños experimentales entre grupos o de caso único, la evaluación de los resultados del tratamiento implica frecuentemente la utilización de cuestionarios. En un enfoque conductual sobre los resultados del tratamiento, se emplean normalmente medidas específicas para evaluar el cambio en un estrecho rango de atención. Sin embargo, el uso de cuestionarios de autoinforme con múltiples centros de atención tiene también su espacio dentro de la evaluación de los resultados de la intervención. De este modo, es importante evaluar los efectos de un programa de intervención no sólo sobre las conductas objetivo principales, sino también en conductas secundarias y en otros aspectos de la vida del paciente. Esta evaluación a múltiples niveles y con múltiples centros de atención puede emplearse para evaluar el impacto de la terapia de conducta sobre numerosas áreas de funcionamiento del paciente. Por ejemplo, el entrenamiento en habilidades sociales que produce una serie de cambios específicos (un aumento del contacto ocular, un volumen de voz más alto, etc.) puede también estar asociado con cambios en la frecuencia de las interacciones interpersonales, el número de personas con las que se asocia el paciente y el ajuste laboral.

Una consideración importante al seleccionar cuestionarios como medidas de los resultados de la terapia implica la sensibilidad de los distintos instrumentos disponibles. La sensibilidad de un instrumento de evaluación se refiere al grado en que puede detectar cambios en la conducta que se está investigando. La selec-

ción de medidas relativamente insensibles puede dar como resultado conclusiones erróneas sobre los efectos de la terapia.

Una buena razón para utilizar cuestionarios con el fin de evaluar los resultados del tratamiento, implica el hallazgo frecuente del fraccionamiento de la respuesta (o desincronía) en la evaluación conductual. Esto se refiere al hecho de que la evaluación del tiempo sistema de respuesta en los problemas de conducta no da frecuentemente como resultado un alto grado de covariación entre los aspectos motores, fisiológicos y cognitivos del funcionamiento del paciente. El uso de un enfoque de medición multimétodo-multiconductas (Cone, 1979) puede ayudar a evaluar las covariaciones entre los tres sistemas de respuesta. Los métodos basados en los cuestionarios pueden emplearse para evaluar los tres contenidos en una serie de conductas. De este modo, los cuestionarios de autoinforme son especialmente apropiados como un método de evaluación para abordar los resultados de la intervención.

VIII. RECOMENDACIONES SOBRE EL EMPLEO DE CUESTIONARIOS DE AUTOINFORME EN LA EVALUACIÓN CONDUCTUAL

En el presente capítulo se ha revisado una serie de cuestiones conceptuales relevantes para la construcción, selección y utilización de cuestionarios en la evaluación conductual. Uno de los principales objetivos ha sido familiarizar al lector con estos temas, de modo que se puedan lograr aplicaciones e interpretaciones más válidas de este método de evaluación conductual. Seguidamente haremos varias sugerencias específicas para aumentar la fiabilidad de los datos obtenidos por los cuestionarios de autoinforme.

- a. Emplear cuestionarios como una fuente, pero no la única, para la obtención de información relativa a la evaluación. Los cuestionarios se utilizan de forma más apropiada como parte de un enfoque de evaluación multimétodo.
- b. Usar cuestionarios con fiabilidad comprobada. Se debería tener en cuenta tanto la estabilidad temporal como la consistencia interna.
- c. Considerar la validez de contenido de los cuestionarios disponibles. Los cuestionarios con especificidad insuficiente o aquellos que evalúan áreas demasiado amplias o demasiado restringidas, o los que no son consistentes con un constructo conductual pueden tener una utilidad limitada.
- d. Considerar la "validez en relación con el criterio" de los cuestionarios disponibles. Es esencial la relación demostrada entre la información proporcionada por los cuestionarios y las conductas emitidas en el ambiente real.
- e. Considerar la sensibilidad de los cuestionarios disponibles. La evaluación sensible ante los aspectos cognitivos, motores y fisiológicos es frecuentemente deseable.
- f. Utilizar cuestionarios que posean normas establecidas para la población

de interés. Es necesario recoger información relativa a las normas para subgrupos.

g. Emplear las puntuaciones de las subescalas además de las puntuaciones totales o resumen. El análisis factorial proporciona un método empírico para obtener estas subescalas.

IX. CONCLUSIONES

Históricamente, muchos terapeutas de conducta han considerado a la información obtenida por los cuestionarios de autoinforme como no válida y poco fiable. Sin embargo, en la última década, una serie de factores ha contribuido a un gradual aumento del uso de los cuestionarios en la evaluación conductual. Este aumento ha permitido que se preste una mayor atención al estudio empírico de distintos instrumentos de autoinforme. Como consecuencia, parece probable que los cuestionarios de autoinforme continúen utilizándose como uno de los enfoques de evaluación más importantes dentro de los programas de evaluación conductual multimétodo.

Una de las principales limitaciones del enfoque de la evaluación por medio de cuestionarios se refiere a la naturaleza indirecta de los datos recogidos (Cone, 1978). Las respuestas del cuestionario implican normalmente marcas con un bolígrafo o con un teclado de ordenador y, como tal, se encuentran lejos de las conductas reales de interés clínico. Sin evidencias que sugieran una relación significativa entre las respuestas a los cuestionarios y las conductas objetivo más directas, los datos obtenidos a partir de los cuestionarios tendrán una utilidad limitada. No es superfluo insistir en la importancia de las evaluaciones relacionadas con la validez del criterio.

REFERENCIAS BIBLIOGRÁFICAS

- Anastasi, A. (1988). *Psychological testing* (6.ª edición). Nueva York: MacMillan.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, Nueva Jersey: Prentice-Hall.
- Bandura, A. (1978). The self-system in reciprocal determinism. *American Psychologist*, 33, 344-358.
- Barlow, D. H. (Comp.) (1981). *Behavioral assessment of adult disorders*. Nueva York: Guilford Press.
- Barrios, B. A. (1988). On the changing nature of behavioral assessment. En A. S. Bellack

- y M. Hersen (Comps.) *Behavioral assessment: a practical handbook* (3.ª edición). Nueva York: Pergamon Press.
- Barrios, B. A. y Hartmann, D. P. (1986). The contributions of traditional assessment: concepts, issues, and methodologies. En R. O. Nelson y S. C. Hayes (Comps.) *Conceptual foundations of behavioral assessment*. Nueva York: Guilford Press.
- Beck, A. T., Ward, C. H., Mendelsohn, M., Mock, J. y Erbugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- Bellack, A. S. y Hersen, M. (Comps.) (1988). *Behavioral assessment: a practical handbook* (3.ª edición). Nueva York: Pergamon Press.
- Bem, D. J. y Allen, A. (1974). On predicting some of the people some of the time: the search for cross-situational consistencies in behavior. *Psychological Review*, 81, 506-520.
- Caballo, V. E. (1993). La multidimensionalidad conductual de las habilidades sociales: propiedades psicométricas de una medida de autoinforme, la EMES-M. *Psicología Conductual*, 1, 221-231.
- Carson, T. P. (1986). Assessment of depression. En A. R. Ciminero, K. S. Calhoun y H. E. Adams (Comps.) *Handbook of behavioral assessment* (2.ª edición). Nueva York: Wiley.
- Cattell, R. B. (1978). *The scientific use of factor analysis*. Nueva York: Plenum Press.
- Ciminero, A. R., Calhoun, K. S. y Adams, H. E. (Comps.) (1986). *Handbook of behavioral assessment* (2.ª edición). Nueva York: Wiley.
- Cone, J. D. (1978). The behavioral assessment grid (BAG): a conceptual framework and a taxonomy. *Behavior Therapy*, 9, 882-888.
- Cone, J. D. (1979). Confounded comparisons in triple response mode assessment research. *Behavioral Assessment*, 1, 85-95.
- Cone, J. D. (1981). Psychometric considerations. En M. Hersen y A. S. Bellack (Comps.) *Behavioral assessment: a practical handbook* (2.ª edición). Nueva York: Pergamon Press.
- Cone, J. D. (1988). Psychometric considerations and the multiple models of behavioral assessment. En A. S. Bellack y M. Hersen (Comps.) *Behavioral assessment: a practical handbook* (3.ª edición). Nueva York: Pergamon Press.
- Cormier, W. H. y Cormier, L. S. (1991). *Interviewing strategies for helpers*. Pacific Grove, California: Brooks/Cole.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Eidelson, R. J. y Epstein, N. (1982). Cognition and relationship maladjustment: development of a measure of dysfunctional relationship beliefs. *Journal of Clinical Psychology*, 50, 715-720.
- Elliot, S. N. (1988). Acceptability of behavioral treatment in educational settings. En J. C. Witt, S. N. Elliot y F. M. Gresham (Comps.) *Handbook of behavior therapy in education*. Nueva York: Plenum Press.
- Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, 37, 1097-1126.
- Epstein, S. (1980). The stability of behavior: II. Implications for research. *American Psychologist*, 35, 790-806.
- Epstein, N. y Eidelson, R. J. (1981). Unrealistic beliefs of clinical couples: their relationship to expectations, goals, and satisfaction. *The American Journal of Family Therapy*, 9, 13-22.