

FIABILIDAD INTERJUECES EN EL PROCESO DE EVALUACIÓN CONDUCTUAL

**Carmen Ramírez Maestre
Rosa Esteve Zarazaga
Antonio Godoy Ávila**

Departamento de Personalidad, Evaluación y Tratamiento Psicológico
Facultad de Psicología. Universidad de Málaga

RESUMEN

El objetivo de este trabajo es la contrastación empírica de los juicios diagnósticos emitidos por una muestra de 114 evaluadores conductuales y de si dichos juicios dependen del problema que se presenta como motivo de consulta y del tipo y número de problemas descritos. Se trata de ver cómo se lleva a cabo el proceso de evaluación conductual, qué metas se proponen y de qué depende la propuesta de unas y no otras. Se envió por correo información escrita de un caso, en términos de comportamientos observables, a un total de 114 psicólogos que manifestaban seguir en su trabajo las orientaciones del modelo teórico cognitivo-conductual . Se enviaron 38 casos del tipo 1 (miedos infantiles), 38 del tipo 2 (problemas de conducta) y 38 del tipo 3 (depresión infantil). De éstos, 57 eran de complejidad baja (tres problemas)

y 57 de complejidad alta (seis problemas). Los resultados indican que, al contrario de lo que esperábamos, presentar una determinada conducta-problema como motivo de consulta hace aumentar su posibilidad de ser elegida como problema principal y como primera meta de tratamiento. Por otro lado, la mayoría de los problemas principales, primeras metas y metas de tratamiento (sin atender al orden temporal en el que se considera que hay que tratarlas), seleccionadas por los evaluadores de nuestra muestra, son factores no descritos de forma explícita en el caso. En cuanto al nivel de complejidad, éste no influye en la fiabilidad interjueces de la selección y ordenación de las metas de tratamiento, obteniéndose niveles muy bajos de acuerdo sea cual sea la naturaleza y complejidad del caso.

Palabras clave: *EVALUACIÓN CONDUCTUAL, PROCESO, FIABILIDAD.*

SUMMARY

One-hundred and fourteen behavioral clinicians provided written replies regarding three brief descriptions of childhood disorders: child fear, conduct disorder, and childhood depression. Each narrative contained either three or six specific presenting problems. Each problem was successively designated across different groups of clinicians as the referral complaint. Presenting a behavior as the referral complaint increased the likelihood of that problem being selected as a main problem and first priority target. Most of the subjects described the treatment goals in terms of the modification of situational factors and the elimination of specific presenting problems. Some of the problems, those "of easy treatment" (enuresis, fears, nailbiting,...) were selected as target behaviors by most of the subjects. The type of case affected the selection of intervention techniques.

Key Words: *BEHAVIORAL ASSESSMENT, PROCESS, RELIABILITY.*

INTRODUCCIÓN

Este trabajo se encuadra dentro del estudio empírico del proceso de evaluación conductual. Más concretamente se incluye entre los estudios empíricos de las garantías de calidad de dicho proceso. La

importancia de este tipo de estudios en el ámbito aplicado radica en varios factores. En primer lugar, aunque parte de la corrección del proceso de evaluación debe quedar asegurada sobre la base de la calidad de las técnicas y procedimientos de recogida de información empleados, esto, por sí sólo, no garantiza la corrección de las decisiones que se tomen al final del proceso total. De cualquier forma, la corrección de los datos proporcionados por las técnicas y procedimientos de evaluación conductual tampoco está respaldada por un número suficiente de trabajos empíricos que garanticen su calidad. Por otro lado, puesto que el trabajo del evaluador se lleva a cabo con muestras pequeñas o con un único sujeto, no contamos con el efecto corrector del manejo de grandes muestras. Esta circunstancia acrecienta la necesidad de garantizar la calidad del proceso. En tercer y último lugar, la mayoría de las veces, el clínico no puede emplear demasiado tiempo, o el cliente no puede costearse un número excesivo de sesiones evaluativas y/o terapéuticas, o no puede demorar la modificación de una situación perjudicial o incluso peligrosa. Por ello, es necesario asegurar que el proceso reúna las garantías de calidad que permitan conseguir los mejores resultados con el mínimo coste.

Por todo esto, resulta de primordial importancia que dicho proceso revista todas las garantías de éxito, para lo cual, parece obligado someterlo a estudio racional y empírico de forma sistemática. Pero, ¿cuáles son las garantías de calidad que ha de cumplir el proceso de evaluación conductual?. Pensamos, junto a no pocos autores (Messick, 1989, 1993; Silva, 1989), que tanto la calidad de las técnicas empleadas en evaluación conductual para la recogida de información, como la calidad del proceso en si, han de ser contrastadas utilizando los criterios psicométricos.

Sin embargo, hasta ahora estos criterios han recibido una escasa atención por parte de los evaluadores conductuales, siendo muy escaso el número de investigaciones empíricas en las que se estudia cómo se lleva a cabo el proceso y en qué medida se encuentran dichos criterios, tanto en los instrumentos de evaluación conductual, como en el proceso en si. Aunque podemos encontrar algunas publicaciones acerca de ciertos aspectos del proceso de evaluación conductual (Beutler y Clarkin, 1990; Hagopian, Fisher, Thompson,

Owen y cols., 1997; Karoly, 1993; Silva, 1993; Turk y Salovey, 1988; Weist, Ollendick y Finney, 1991), no encontramos trabajos empíricos que contrasten la calidad del mismo. Si, como creemos, se han de aplicar los criterios psicométricos de calidad a la evaluación conductual, y más concretamente al proceso, es el momento de ver cómo se ha de realizar el mismo para determinar a continuación, las garantías de calidad que presenta.

Algunos autores señalan la importancia de contar con un proceso normativo, en el que se estipulen los distintos pasos y actividades a seguir por el evaluador, para maximizar así el rigor científico de la evaluación psicológica (Westmeyer y Hageböck, 1992). Lo que aconsejan la mayoría de los evaluadores conductuales al respecto está basado únicamente en la «experiencia» y no en la investigación, y se podría resumir del siguiente modo: El cliente acude a consulta con una serie de quejas y demandas que el terapeuta habrá de traducir en términos conductuales específicos, identificando así las conductas problema. La siguiente fase la constituye el análisis de las variables controladoras de estas conductas. A continuación, cliente y terapeuta decidirán los objetivos o metas últimas de tratamiento, es decir, el estado B al que queremos llegar a partir del estado actual A del sujeto. Basándonos en este objetivo final se seleccionarán las conductas meta o conductas objetivo, que hacen referencia a las variables concretas de la conducta o del contexto en la que ésta sucede y sobre las que se focalizarán los efectos del tratamiento. La modificación de estas conductas deberá producir un efecto que constituye lo que se ha denominado objetivo meta o meta última de tratamiento. Para guiar en la selección de estas conductas-meta se han propuesto una serie de criterios empíricos y criterios basados en la experiencia de los autores. Los criterios empíricos propuestos, por otra parte, son más bien procedimientos de investigación básica que proporcionan más información y no consejos basados en la investigación empírica de la práctica clínica ordinaria. Por otro lado, en esta fase se incluye la selección de la conducta meta prioritaria. Para llevar a cabo esta elección también se han propuesto una serie de características de las conductas problema que ayudan al evaluador a tomar una decisión adecuada al respecto. Todos estos consejos de nuevo se basan en

la experiencia clínica de los distintos autores. Una vez que sabemos sobre lo que intervenir deberemos decidir cómo hacerlo, es decir habremos de seleccionar las técnicas de tratamiento que deberemos utilizar para la consecución de las metas propuestas. Por último se evaluará el tratamiento durante la aplicación del mismo y tras su finalización, realizándose una valoración de los resultados obtenidos.

Posiblemente, los tres grandes trabajos acerca de la calidad del proceso de evaluación conductual son los llevados a cabo por Hay, Hay, Angle y Nelson (1979), Wilson y Evans (1983) y Felton y Nelson (1984), cuyos resultados muestran un acuerdo moderado o bajo entre los evaluadores conductuales en las decisiones tomadas durante el proceso de evaluación. En definitiva, todos estos trabajos apuntan la necesidad de llevar a cabo más estudios de este tipo, intentando identificar cuáles pueden ser las razones de la falta de acuerdo en los distintos momentos de la evaluación conductual.

OBJETIVO E HIPÓTESIS

El objetivo de este trabajo es la contrastación empírica de los juicios diagnósticos emitidos por los evaluadores conductuales españoles de nuestra muestra y de si dichos juicios dependen del problema que se presenta como motivo de consulta y del tipo y número de problemas descritos. Se trata de ver cómo se lleva a cabo el proceso de evaluación conductual, qué metas se proponen y de qué depende la propuesta de unas y no otras.

Como ya se ha comentado, la cantidad de estudios empíricos e incluso teóricos acerca de la calidad del proceso de evaluación conductual es muy escasa. Por ello, la mayoría de las hipótesis que pasamos a describir a continuación se derivan de lo que se espera que realice un evaluador conductual durante el proceso, según los consejos normativos del modelo teórico al que pertenecen, así como de los resultados obtenidos por Wilson y Evans (1983) en el trabajo empírico mencionado en líneas anteriores. Dichas hipótesis son las siguientes:

Sobre el Motivo de Consulta

Los primeros estudios acerca de la fiabilidad de los diagnósticos psiquiátricos identificaron el motivo de consulta como uno de los factores que más influye en la identificación de los problemas principales del cliente. De hecho, los evaluadores tienden a ignorar otros problemas a favor de dicha queja (Arnhoff, 1954; Ward, Beck, Mendelson, Mock, and Erbaugh, 1962). En terapia de conducta, definir los problemas con relación a la preocupación más inmediata del cliente ha sido un regla implícita (Kanfer and Saslow, 1969; Mash and Terdal, 1979). Sin embargo, a menudo, la interpretación que hace el cliente de sus problemas no es la mejor, y el clínico debe hacer uso de su experiencia y conocimientos para seleccionar el principal problema del sujeto y la primera meta de tratamiento (Haynes, 1979; O'Leary and Wilson, 1987). Basándonos en los resultados obtenidos por Wilson y Evans (1983) en su estudio empírico, así como en los trabajos de otros autores, (Elstein, Shulman y Sprafka, 1978; Hawkins, 1986; Haynes, 1979; Kanfer y Saslow, 1969; Mash y Terdal, 1979; O'Leary y Wilson, 1987), se formulan varias hipótesis al respecto: (1) Creemos que el presentar uno de los problemas descritos en el caso, como motivo de consulta por parte del cliente, aumentará la probabilidad de que éste sea elegido como meta de tratamiento, siendo así seleccionado más consistentemente por los encuestados como objetivo instrumental. (2) Por otro lado, partiendo de los resultados de los trabajos de Hay y colaboradores (1979) y Wilson y Evans (1983), según los cuales la conceptualización que hacen del caso los evaluadores conductuales es distinta de la propuesta de tratamiento, pensamos que la elección del problema principal que aqueja al cliente y (3) de la primera meta de tratamiento será independiente del comportamiento que constituya la demanda o queja del mismo. Es decir, creemos que el clínico, para escoger el problema principal del cliente y la primera meta tendrá más en cuenta su experiencia y conocimientos al respecto que la demanda.

Sobre la complejidad

Basándonos en la idea, defendida por un buen número de autores e investigaciones al respecto (véase Elstein y colaboradores, 1978),

según la cual, los evaluadores tienen dificultades para trabajar con mucha información, suponemos que (4) existe una relación inversa entre fiabilidad y complejidad, esto es, a mayor nivel de complejidad (operativizado como número de conductas-problema presentes en el caso), menor acuerdo entre jueces, aumentando éste al disminuir el número de problemas del caso. Esta hipótesis es contraria a los resultados obtenidos por Wilson y Evans (1983), sin embargo, pensamos que éstos pudieron deberse al hecho de que un porcentaje considerable de evaluadores (32%), seleccionaron como meta de tratamiento ciertos factores inferidos, no presentes en el caso. De esta forma, el número de conductas problema no influyó en la selección de las metas de tratamiento, minimizándose así la distinción entre casos en función de su complejidad. Es evidente que si los encuestados seleccionan las metas de tratamiento entre las conductas-problema presentadas (como de hecho cabría esperar en una muestra de terapeutas conductuales), la probabilidad de acuerdo es doblemente mayor al escoger entre tres tipos de problemas que si la selección se realiza entre un grupo de seis.

Sobre el tipo de caso

Parece lógico que la decisión de los evaluadores en el proceso se vea influida por sus conocimientos de psicopatología. Por ejemplo, los problemas más comunes entre los niños, que no suelen perdurar más allá de la primera infancia podrían ser considerados como menos susceptibles de ser tratados (Miller, Barrett and Hampe, 1974; Schechtman, 1970) en comparación con otros problemas de peor pronóstico (Robins, 1979). Por otro lado, ciertos autores señalan que se suele seleccionar como objetivo de tratamiento aquellas conductas cuya modificación parece más fácil (Haynes, 1979; Nelson and Hayes, 1986; O'Leary, 1979; Schulte, 1974). Así pues, de nuevo basándonos en estos estudios teóricos y en los resultados del trabajo de Wilson y Evans (1983), se hipotetiza que la naturaleza de los casos expuestos influirá en varios aspectos: (5) En la decisión de in-

tervenir, de forma que entre aquellos encuestados que recibieron un caso de «miedos infantiles» habrá un porcentaje mayor de sujetos que decidan no intervenir que entre los encuestados que recibieron los casos de «trastornos de conducta» o «depresión infantil». Suponemos que esto es así porque la literatura psicopatológica nos dice que los miedos infantiles suelen remitir con la edad, causando un trastorno menor que los otros dos tipos de problemas. (6) De igual forma, basándonos en ciertos autores, según los cuales se suele seleccionar como objetivo de tratamiento aquellas conductas cuya modificación parece más fácil (Haynes, 1979; Nelson y Hayes, 1986; O'Leary, 1979 y Schulte, 1974) se espera que aquellas conductas-problema para cuya modificación se disponga de técnicas de intervención estandarizadas y de alto porcentaje de éxito (ej. enuresis, tartamudeo y onicofagia), serán seleccionadas con mayor probabilidad como metas de tratamiento.

Variables

Las variables de este estudio se dividen en dos grandes grupos: variables independientes y dependientes.

- *Variables Independientes*. Las variables independientes son tres y aparecen a través de la manipulación de la información que se presentó a cada uno de los sujetos de nuestra muestra:

1. Motivo de Consulta. Se trata de la demanda o queja con que se presenta el cliente a consulta. Cada uno de los problemas descritos en los tres diferentes casos, se presentó como motivo de consulta, al comienzo del informe, dependiendo de la condición experimental. Como veremos más adelante, contamos con tres tipos de casos distintos: miedos infantiles, trastornos de conducta y depresión infantil. Cada uno de ellos incluye seis problemas. A su vez, cada problema fue presentado como motivo de consulta en cada caso según la condición experimental. Por lo tanto, contamos con dieciocho motivos de consulta posibles: seis problemas por cada uno de los tres diferentes casos.

2. Complejidad del Caso. Esta variable se definirá por el número de problemas presentes en el caso. Tratamos con dos niveles de

complejidad: alta y baja, siendo de alta complejidad los casos con seis problemas y de baja los de tres. Por lo tanto, cada uno de los tres casos adquiere dos formas distintas en función de su complejidad, pudiendo aparecer con seis problemas o sólo con tres de éstos.

3. Tipo de Caso. Esta variable se refiere a la naturaleza de los distintos problemas presentados. Como ya hemos comentado, el tipo o naturaleza de los problemas varía a través de la presentación de tres casos distintos: «miedos infantiles», «problemas de conducta», y «depresión infantil». Se trata de tres casos infantiles ficticios, creados por el equipo de investigación, cuya sintomatología ha sido derivada del DSM-III (APA, 1980), de la literatura al respecto (Cantwell and Carlson, 1987; Del Barrio, 1984; Lefkowitz and Burton, 1978; Moreno and Valero, 1987; Pelechano, 1981; Romero, 1988; Romero et al., 1989) y de los casos utilizados por Wilson y Evans (1983). Cada caso describe uno de los tres principales tipos de problemas infantiles identificado en estudios sobre análisis factorial (Achenbach y Edelbrock, 1978). Todos los casos fueron corregidos por tres jueces independientes, psicólogos clínicos con una extensa experiencia en terapia. Los informes incluyen una breve descripción del desarrollo del niño y de su ambiente familiar. Además, entre los problemas actuales que presenta el sujeto se añaden algunos ítems de relleno tales como hechos ocurridos en el pasado o problemas de menor dificultad que proporcionan una mayor verosimilitud al caso expuesto.

- *Variables Dependientes.* Sobre los casos enviados se formularon cuatro preguntas, considerando los juicios diagnósticos emitidos por los sujetos las variables dependientes. Estas son:

1. Necesidad de tratamiento. Esta variable queda definida por la respuesta del encuestado a la primera pregunta del cuestionario presentado, en la cual se pedía al evaluador su juicio acerca de la necesidad de intervenir en el caso expuesto.

2. Problema principal. Será la respuesta del encuestado a la segunda pregunta del cuestionario, en la que se pedía al evaluador que seleccionara el problema más importante o destacado del sujeto del caso presentado.

3. Metas de tratamiento. Esta variable se define como el juicio diagnóstico del sujeto acerca de las conductas, situaciones o grupos de conducta que se han de seleccionar como objetivos instrumentales de tratamiento.

4. Primera meta de tratamiento. Como la variable anterior, ésta quedará definida por la respuesta del encuestado a la tercera pregunta del cuestionario, en la cual se solicitaba el orden temporal de consecución de cada meta de tratamiento seleccionada. Así, aquella conducta, situación o grupo de conductas seleccionado como primer objetivo instrumental a conseguir sería la variable a la que nos referimos.

5. Técnicas de tratamiento. Será la respuesta del encuestado a la cuarta pregunta del cuestionario, en la que se pedía al evaluador que seleccionara las técnicas de intervención que utilizaría para conseguir cada una de las metas propuestas.

Para intentar aclarar las diferencias entre estas variables proponemos un ejemplo: uno de los casos enviados a los evaluadores de nuestra muestra hace referencia a un niño que presenta problemas de sueño, de interacción con sus iguales y onicofagia. El padre del niño lo lleva a consulta para solucionar la onicofagia (*motivo de consulta*), sin embargo, el evaluador piensa que el *problema principal* lo constituye la deficiente interacción social del sujeto, aunque lo primero que querrá conseguir el terapeuta es que el niño duerma bien (*primera meta de tratamiento*), para incidir posteriormente en el resto de conductas-problema. En este ejemplo, el clínico no propone como primera meta de tratamiento la intervención sobre el motivo de consulta ni sobre el problema principal, es decir, ni sobre lo que el cliente considera más importante ni sobre lo que el mismo evaluador considera más importante, sino sobre lo que éste piensa que es más urgente o cuya solución ayudará a conseguir las siguientes metas de tratamiento. Por supuesto, estas variables no tienen por qué hacer referencia a problemas distintos. De hecho, cliente y terapeuta pueden estar de acuerdo en cuál es el problema más importante, coincidiendo entonces *motivo de consulta* y *problema principal*. También es posible que el clínico piense que la *primera meta de tratamiento* deba coincidir con *el motivo de consulta*, aunque considere que *el principal problema* es otro.

MÉTODO

Sujetos

La muestra está constituida por 114 psicólogos españoles que manifiestan seguir una orientación teórica cognitivo-conductual en su trabajo. El 51% son mujeres frente al 49% de hombres; la edad media es de 33 años, oscilando entre 26 y 57, con una desviación típica en torno a 7; la actividad profesional principal de los encuestados es la clínica (68%), dedicándose el resto a la enseñanza, orientación, investigación y otras. Por último, los sujetos de la muestra presentan una media de 7 años y medio de experiencia clínica.

Diseño

Se envió por correo información escrita de un caso, en términos de comportamientos observables, a un total de 311 psicólogos. De los cuestionarios que nos fueron devueltos se seleccionaron 114 que cumplían los requisitos del presente estudio (se igualaron los grupos para cada condición siempre con clínicos que manifestaban seguir las orientaciones del modelo cognitivo-conductual). Se seleccionaron en fin 38 casos del tipo 1 (miedos infantiles), 38 del tipo 2 (problemas de conducta) y 38 del tipo 3 (depresión infantil). De éstos, 57 eran de complejidad baja (tres problemas) y 57 de complejidad alta (seis problemas).

Los 38 casos del tipo 1 (miedos infantiles), contenían los siguientes problemas: enuresis, celos, miedo a la oscuridad, miedo a los adultos desconocidos, miedo a los perros y problemas con las comidas. Los 38 del tipo 2 (problemas de conducta), presentaban destrucción, desobediencia, mentiras, cambios de humor, agresividad y tartamudeo. Por último, los 38 del tipo 3 (depresión infantil), presentaban problemas de atención, problemas de interacción con sus iguales, problemas de sueño, dolores y vómitos, onicofagia y apatía.

Junto al caso se hacen unas preguntas acerca de ciertos datos biográficos tales como sexo, edad, años de experiencia profesional, lugar de empleo, categoría profesional, orientación teórica y activi-

dad profesional y se formulan al clínico las cuatro cuestiones que recogen los juicios diagnósticos que suponen las variables dependientes. Este cuestionario es anónimo y las preguntas se formularon de forma abierta para no condicionar las respuestas. Éstas se analizaron de acuerdo con un sistema de categorías desarrollado a partir de los datos. La fiabilidad de dicho sistema de categorías se determinó por la calificación independiente de las mismas por parte de dos jueces. La fiabilidad interjueces de las categorías se calculó mediante el índice *Kappa* de fiabilidad interjueces (Cohen, 1966; Suen, Ary y Ary, 1986), aplicándose a cada categoría. El acuerdo es bastante alto en todos los casos, oscilando las medias entre .84 y .99 en las distintas categorías.

RESULTADOS

Sobre los resultados se han realizado varios tipos de análisis estadísticos: porcentajes, razones de probabilidades, la prueba «*Chi-cuadrado*» y el coeficiente de concordancia *W de Kendall*. La prueba *Chi-cuadrado* se ha utilizado para determinar la existencia de diferencias en las frecuencias de los distintos grupos experimentales respecto a los juicios diagnósticos emitidos. El coeficiente *W de Kendall* se ha calculado en cada grupo para analizar el grado de concordancia en la ordenación de las metas de tratamiento seleccionadas. *W* puede variar de cero a uno y puede interpretarse como el resto de los coeficientes de correlación. Así, el cero representará la ausencia de correlación mientras que el uno representa la correlación perfecta.

Sobre el motivo de consulta

En la tabla I presentamos el número y porcentaje de ocasiones en las que se seleccionan las conductas-problema como problema principal, dependiendo de si éstas han sido o no presentadas como motivo de consulta. Sobre estas cifras se calculó la razón de probabilidad que nos indicará si el hecho de presentar una conducta-problema como motivo de consulta altera su probabilidad de ser seleccionada como problema principal en relación a cuando no se

Tabla 1.- Influencia del motivo de consulta sobre la elección del problema principal, la primera meta de tratamiento y las metas de intervención. Número y porcentaje de respuestas

	Si la conducta-problema es presentada como motivo de consulta	Si la conducta problema no es presentada como motivo de consulta	TOTAL
conducta-problema seleccionada como problema principal	N = 15 % = 3.1	N = 7 % = 1.4	N = 22 % = 4.5
conducta-problema no seleccionada como problema principal	N = 99 % = 20.2	N = 368 % = 75.3	N = 467 % = 95.5
TOTAL	N = 114 % = 23.3	N = 375 % = 76.7	N = 489 % = 100
RAZÓN DE PROBABILIDAD	$\frac{15 / 114}{7 / 375} = 7.04$		
conducta-problema seleccionada como primera meta	N = 12 % = 2.45	N = 16 % = 3.27	N = 28 % = 5.72
conducta-problema no seleccionada como primera meta	N = 102 % = 20.86	N = 359 % = 73.42	N = 461 % = 94.28
TOTAL	N = 114 % = 23.3	N = 375 % = 76.7	N = 489 % = 100
RAZÓN DE PROBABILIDAD	$\frac{12 / 114}{16 / 375} = 2.467$		
conducta-problema seleccionada como meta de tratamiento	N = 18 % = 3.7	N = 89 % = 18.2	N = 107 % = 21.9
conducta-problema no seleccionada como meta de tratamiento	N = 96 % = 19.6	N = 286 % = 58.5	N = 382 % = 78.1
TOTAL	N = 114 % = 23.3	N = 375 % = 76.7	N = 489 % = 100
RAZÓN DE PROBABILIDAD	$\frac{18 / 114}{89 / 375} = 0.66$		

la presenta como tal. Este análisis nos indica que es 7 veces más frecuente seleccionar una conducta como problema principal cuando ésta constituye el motivo de consulta que cuando no lo es. Por lo tanto, la hipótesis propuesta parece quedar refutada en cuanto que, al contrario de lo que esperábamos, es más probable que una conducta sea seleccionada como problema principal si ésta constituye el motivo de consulta.

Por otro lado, es importante señalar lo que mas frecuentemente eligen los clínicos de esta muestra como problema principal. Con el fin de saber qué elegían, realizamos un análisis exploratorio que nos indicara el número de sujetos que seleccionan cada categoría de problema principal, dependiendo de la conducta problemática que se les haya presentado como motivo de consulta. A modo descriptivo podemos observar que de los 155 juicios emitidos sobre cuál es el problema principal (encontramos más juicios que sujetos ya que algunos clínicos señalaron varios problemas principales), el 37% se refieren a expresiones globales (por ejemplo, problemas de conducta), el 29% a factores situacionales inferidos (por ejemplo, ambiente desestructurado), el 14% a conductas concretas (por ejemplo, eneuresis), y el 12% a etiquetas diagnósticas (por ejemplo, fobias). Menos del 6% considera que la información proporcionada es insuficiente y, por tanto, suspende su diagnóstico.

Así pues, lo que se esperaba que fuera la categoría más seleccionada (elegir como problema principal conductas específicas descritas en el caso) ocupa un tercer lugar, seguida muy de cerca por las etiquetas diagnósticas. Por otro lado, las expresiones globales que incluyen varias conductas descritas en el caso (la categoría más frecuentemente seleccionada) son muy utilizadas por nuestra muestra, aun cuando probablemente sean peores que las etiquetas diagnósticas a las que parecen querer sustituir ya que les falta precisión descriptiva y no permite ni pronóstico ni indicación de tratamiento. En resumen, presentar una conducta como motivo de consulta favorece el que se la elija como el problema principal del sujeto. No obstante, lo más usual es que el problema principal se exprese en términos distintos a los utilizados en la descripción del caso. Especialmente frecuente es, en este sentido, el uso de categorías y expresiones globales, generales y poco precisas.

De nuevo en la tabla I se recoge la frecuencia y porcentaje de ocasiones en la que se selecciona como primera meta de tratamiento cada una de las conductas-problema presentadas en el caso expuesto, dependiendo de si éstas han sido presentadas como motivo de consulta o no. La razón de probabilidad nos indica que es 2 veces más frecuente seleccionar una determinada conducta como primera meta cuando aquélla constituye el motivo de consulta que cuando no lo es. Por lo tanto, nuestra hipótesis al respecto queda refutada ya que el presentar una conducta como motivo de consulta dobla la probabilidad de que se la escoja como primera meta de tratamiento. Sin embargo, la mayoría de las primeras metas de tratamiento que proponen los clínicos son categorías o factores que no aparecen explícitamente en la descripción del caso, como ya hemos visto que ocurre también en la elección del problema principal.

Así, de las 118 primeras metas de tratamiento seleccionadas, el 44% se refieren a la modificación de factores situacionales, el 19% propone estrategias de evaluación, el 13% términos globales de cambio, el 11% hace referencia a la intervención, definida en términos «negativos» sobre problemas específicos descritos en el caso (por ejemplo, eliminar la conducta problema), mientras el 7% propone la intervención «positiva» sobre los mismos (por ejemplo, desarrollar una conducta adaptativa que sustituya a la conducta problema) y el 2% estrategias específicas de tratamiento. Por lo tanto, la modificación de factores situacionales es seleccionada mayoritariamente como primera meta de tratamiento. Por otro lado, el hecho de que un 19% se refiera a estrategias de evaluación para expresar la primera meta puede que se deba a que estos evaluadores requieren más información antes de proponer las metas de tratamiento, aunque nadie señaló que la información proporcionada no fuera suficiente.

Por último, en la tabla I se presenta la frecuencia y porcentaje de ocasiones en las que se seleccionan como metas de tratamiento, cada uno de los problemas presentados en el caso expuesto. Se compararán las diferencias entre aquéllos que recibieron dichos problemas como motivo de consulta y los que no. El resultado del análisis nos indica que presentar una conducta como queja o demanda del cliente no hace que se la elija más como meta de tratamiento, quedando de nuevo refutada nuestra hipótesis.

Son muchas las ocasiones en las que los sujetos de nuestra muestra seleccionan como meta de tratamiento factores o variables que no aparecen como conductas-problema (sean o no motivos de consulta) en el caso. Se puede observar que de las 334 metas de tratamiento propuestas por los sujetos de nuestra muestra, el 25% se refieren a la modificación de factores situacionales, el 21% a la intervención, definida en términos «negativos», sobre problemas específicos descritos en el caso, el 20% proponen términos globales de cambio, el 12% estrategias específicas de tratamiento, el 11% la intervención «positiva» sobre problemas específicos descritos en el caso y el 10% propone estrategias de evaluación. Si bien nadie señaló que la información proporcionada no fuera suficiente para la elección de las metas de tratamiento, el hecho de que hayan propuesto estrategias de recogida de información (entrevistas, observación, etc.) puede hacer pensar que no consideran suficientes los datos de que disponen.

Sobre el nivel de complejidad

Puesto que el análisis de la fiabilidad interjueces, en la ordenación de las metas de tratamiento, se realizó simultáneamente con las variables independientes *nivel de complejidad y tipo de caso*, en las tablas que presentamos a continuación se contrasta la hipótesis sobre la influencia de la complejidad a la vez que se obtiene información adicional respecto al tipo de caso. En la tabla 2 se recoge la fiabilidad entre sujetos en la ordenación de las metas de tratamiento. Para hallarla se ha utilizado el coeficiente de concordancia W de Kendall. Podemos observar que el acuerdo entre los jueces es muy bajo en todos los casos y niveles de complejidad. El hecho de que sean significativos proviene del alto número de respuestas analizadas.

Sobre el tipo de caso

Existe un claro acuerdo en la necesidad de intervenir en los casos presentados, sea cual sea la naturaleza de los mismos. Por lo tanto,

Tabla 2.- Fiabilidad interjueces en la ordenación de las metas de tratamiento. Coeficientes de acuerdo W de Kendall

Casos	W _s		
	Baja complejidad (3 problemas)	Baja complejidad (3 problemas)	Alta complejidad (6 problemas)
Caso 1 (miedos infantiles)	.31 **	.18 **	.35 **
Caso 2 (problemas de conducta)	.14 **	.17 **	.13 **
Caso 3 (depresión infantil)	.29 **	.23 **	.22 **

** p < .0001

de nuevo se refuta nuestra hipótesis al respecto ya que el tipo de caso no influye en este juicio diagnóstico.

En la tabla 3 queda reflejado el porcentaje de ocasiones en las que se selecciona como meta de tratamiento, cada uno de los problemas específicos descritos en el caso. La importancia de las diferentes metas vendrá determinada por el número de sujetos que, habiéndole sido presentada esa conducta, la seleccionan, independientemente del orden propuesto. Así pues, podemos observar que en el caso tipo tres, la onicofagia es elegida como meta de tratamiento en 14 ocasiones, seguida por los problemas de interacción, los cuales fueron seleccionados como metas de tratamiento en 8 ocasiones. Si atendemos a las conductas-problema que presenta este caso (problemas de atención, de interacción, alteraciones del sueño, dolores y vómitos, onicofagia y apatía), la elección más frecuente de la onicofagia como meta de tratamiento se adecua a la hipótesis formulada, según la cual, las conductas de «fácil» tratamiento son las más frecuentemente seleccionadas como metas.

En el caso tipo dos, agresividad y tartamudeo son seleccionadas prácticamente el mismo número de ocasiones (13 y 14 respectivamente) y, si bien el tartamudeo podría encontrarse entre aquellos problemas para los que contamos con tratamientos estandarizados y con buenos porcentajes de éxito, la intervención sobre las conductas agresivas resulta menos clara. Por otro lado, en el caso tipo uno, aunque el número de ocasiones en las que se seleccionan la mayoría de las conductas es similar (siendo menos seleccionadas «celos» y «miedo a la oscuridad»), «eneuresis», «miedo a adultos desconocidos» y «miedo a los animales» se encuentran entre las más escogidas como metas de tratamiento.

De esta forma, podemos decir que conductas-problemas para las que contamos con procedimientos terapéuticos de aplicación específica, estandarizados y muy utilizados en modificación de conducta (eneuresis, miedo a los animales, tartamudeo y onicofagia), se encuentran entre las más seleccionadas por nuestros evaluadores como comportamientos objetivo de tratamiento.

Por último, realizamos un análisis descriptivo exploratorio que proporciona datos acerca de qué tipos de técnicas de tratamiento son seleccionadas para intervenir sobre los casos descritos. Los

Tabla 3.- Porcentaje y número de ocasiones en las que cada problema es seleccionado como meta de tratamiento

Problemas	Total	
	N	%
Caso 1 (miedos infantiles)		
1. enuresis	7	21.9
2. celos	2	6.2
3. miedo a la oscuridad	3	9.4
4. miedo a adultos desconocidos	8	25
5. miedo a los animales	6	18.7
6. poco apetito	6	18.7
Total	32	100
	N	%
Caso 2 (problemas de conducta)		
1. destrucción	4	10.5
2. desobediencia	3	7.9
3. mentiras	4	10.5
4. cambios de humor	0	0
5. agresividad	13	34.2
6. tartamudeo	14	36.8
Total	38	100
	N	%
Caso 3 (depresión infantil)		
1. falta de atención	2	2.7
2. problemas de interacción	8	21.6
3. problemas de sueño	6	16.2
4. dolor y vómitos	4	10.8
5. onicofagia	14	37.8
6. apatía	3	10.8
Total	37	100

Tabla 4.- Porcentaje y número de sujetos que seleccionan cada categoría de técnicas de tratamiento

Categorías	Caso 1 (miedos)	Caso 2 (problemas de conducta)	Caso 3 (depresión)	Chi cuadrado
1. técnicas de condicionamiento operante (e.g. extinción)	N = 24 % = 82.7	35 74.5	31 81.6	.99
2. técnicas de reducción de ansiedad (e.g. desensibilización sistemática)	N = 23 % = 79.3	18 38.3	24 63.1	13.57 ****
3. técnicas cognitivas y de autocontrol (e.g. solución de problemas)	N = 5 % = 17.2	15 31.4	22 57.9	11.06 ***
4. juego de roles	N = 6 % = 20.7	6 12.8	9 23.7	1.60
5. modelado	N = 4 % = 13.8	11 23.4	6 15.8	1.46
6. entrenamiento en habilidades sociales	N = 7 % = 24.1	6 12.8	19 50	13.43 ****
7. entrenamiento a padres	N = 18 % = 62.1	13 27.6	18 47.4	9.38 **
8. terapia de parejas	N = 1 % = 3.4	7 14.9	11 28.9	7.21 *
9. términos inespecíficos	N = 13 % = 44.8	24 51.1	22 57.9	.65
10. estrategias de evaluación	N = 14 % = 48.3	14 29.8	17 44.7	3.13
11. otros	N = 8 % = 27.6	9 19.4	3 7.9	5.06 +

+ $p < .07$ (tendencia estadística)

* $p < .05$ *** $p < .005$

** $p < .01$ **** $p < .001$

resultados quedan reflejados en la tabla 4. En ella podemos observar que las técnicas de reducción de ansiedad son seleccionadas más a menudo para el caso 1 (miedos infantiles), y las técnicas cognitivas para el caso 3 (depresión infantil).

CONCLUSIONES

Seguramente, llama mucho la atención el hecho de que los resultados obtenidos en este trabajo no apoyen la mayoría de las hipótesis formuladas. Puesto que gran parte de dichas hipótesis se basa en lo que se supone que debe hacer un psicólogo clínico que siga los consejos normativos propuestos por el modelo cognitivo-conductual, estos datos revelan importantes diferencias entre el proceso de evaluación conductual propuesto por el modelo teórico y el proceso evaluativo real llevado a cabo en la práctica clínica.

La práctica totalidad de los encuestados coinciden en la necesidad de intervenir en todos los casos presentados, sea cual sea su naturaleza, complejidad y motivo de consulta. Este acuerdo, puede deberse a una creencia general de que es necesario tratar todos, o prácticamente todos, los problemas que llegan a consulta. Por lo tanto, estos evaluadores no tienen en cuenta que, de acuerdo con la literatura psicopatológica (Miller, Barrett y Hampe, 1974; Robins, 1979; Schechtman, 1970), los miedos descritos son relativamente comunes entre los niños y tienden a remitir tras la primera infancia.

Presentar un problema como motivo de consulta afecta a la percepción que hace el evaluador de dicha conducta, de forma que será seleccionada como problema principal y primera meta de tratamiento más a menudo si constituye motivo de consulta que si es incorporada al caso como un problema más. Así pues, el clínico parece valorar el caso de igual forma que el cliente, independientemente de otras consideraciones presentes en los estudios teóricos y empíricos de dichos tipos de caso.

Cuando se pide al evaluador que seleccione el principal problema de cada caso, éstos tienden a utilizar expresiones globales que incluyen varias de las conductas problema presentes. Sin embargo, como evaluadores conductuales, se espera que seleccionen una de

entre las diferentes conductas-problema descritas en el caso. Parece que estas expresiones globales tienden a sustituir a las tan denostadas etiquetas diagnósticas. Sin embargo, el uso de expresiones globales es menos aconsejable que el de las etiquetas diagnósticas, en tanto que la definición de éstas queda delimitada en los manuales diagnósticos mientras que el significado de aquellas depende en gran medida del evaluador que las utiliza. De cualquier forma, el hecho de que la construcción de los casos se haya basado en manuales psicodiagnósticos tales como el DSM-III, de forma que el conjunto de los problemas compone una categoría diagnóstica típica, probablemente ha influido en la utilización de expresiones globales, e incluso de alguna etiqueta diagnóstica.

Otro de los juicios diagnósticos requeridos es la especificación de las metas de tratamiento. Las respuestas se clasificaron en categorías entre las cuales la más frecuente es aquella que se refiere a la eliminación de conductas-problema descritas en el caso. Podemos observar por lo tanto que, a pesar de que en la literatura de evaluación conductual se aconseja de forma insistente plantear el tratamiento bajo un enfoque constructivo, el enfoque teórico «eliminativo» es mayoritario entre los evaluadores conductuales. Así, el 21.2% de nuestra muestra propone la eliminación de conductas-problema como meta de tratamiento, mientras un 10.8% propone la implantación de conductas positivas alternativas. De nuevo encontramos diferencias entre la teoría planteada y la práctica clínica diaria del evaluador conductual.

Las conductas seleccionadas con mayor frecuencia como metas de tratamiento son aquéllas para las que contamos en modificación de conducta, con tratamientos estandarizados. Esto nos hace pensar que los evaluadores conductuales no realizan un análisis funcional del caso si pueden evitarlo. Es decir, seleccionan muy frecuentemente como metas de tratamiento aquéllas conductas para las que existe, en evaluación conductual, «recetas», independientemente de que dichos remedios son los más adecuados para el cliente en cuestión. ¿Acaso no se sigue de esta forma el tan criticado modelo médico?. Seguramente habremos de preguntar, como ya lo hace Schulte (1989), acerca de la utilidad de tratamiento de la evaluación conductual frente a modelos categoriales.

Por último, puesto que el grado de acuerdo entre los evaluadores en la selección de las metas de tratamiento y en el orden propuesto de las mismas es bastante bajo en todos los casos y niveles de complejidad, parece que aunque cuenten con la misma información, el uso que hacen de la misma así como las conclusiones evaluativas a las que llegan distan mucho de coincidir. Así pues, tanto que se ha criticado en evaluación conductual, los bajos niveles de acuerdo conseguidos en los modelos de evaluación tradicionales, ahora, mediante los estudios acerca de la calidad de la evaluación conductual comprobamos que la concordancia entre evaluadores, en distintos momentos del proceso, es bastante baja. Por lo tanto, la creencia ingenuamente realista que imperó durante mucho tiempo en el enfoque conductual, según la cual la evaluación conductual, sólo por ser conductual ya presentaba garantías de calidad, no se mantiene hoy día. Creemos pues que la investigación empírica de la calidad del proceso, así como de las variables que influyen en la misma es muy importante y puede ayudarnos a salir airoso de la actual crisis.

En resumen, los evaluadores conductuales de este estudio: (1) intervienen sobre todos los casos que llegan a consulta, (2) definen el caso del mismo modo que el cliente, (3) seleccionan como objetivo de tratamiento sobre todo aquellos problemas para los que existe una técnica de tratamiento conductual tradicional, probablemente movidos por una búsqueda de “recetas terapéuticas”, (4) se mueven en un enfoque “eliminativo” frente al “constructivo” y (4) consiguen bajos niveles de acuerdo en la selección de las metas de tratamiento. Así pues, como ya se ha señalado, parece haber un abismo entre el modelo teórico y la práctica profesional y, si esto es así, deberíamos cuestionarnos si la forma en la que enseñamos a los futuros clínicos es la apropiada. Creemos necesaria la continuación de estudios empíricos en esta línea que nos permitan conocer el “quehacer” diario del psicólogo clínico, así como la fiabilidad de sus decisiones. De esta forma se podría contribuir a la creación de un modelo teórico coherente que guíe al evaluador en la toma de decisiones durante el proceso evaluativo.

BIBLIOGRAFÍA

- Achenbach, T. M. y Edelbrock, C. S.** (1978). The Classification of Child Psychopathology: A Review and Analysis of Empirical Efforts. *Psychological Bulletin*, 85, 1275-1301.
- APA** (1980). *Diagnostic and Statistical Manual of Mental Disorders (III)*. Washington: APA.
- Arnhoff, F. N.** (1954). Some Factors Influencing the Unreliability of Clinical Judgments. *Journal of Clinical Psychology*, 10, 272-275.
- Beutler, L.E. y Clarkin, J.F.** (1990). *Systematic treatment selection: Toward targeted therapeutic interventions*. Nueva York: Brunner/Mazel. Usement in Data-Based Intervention Studies. *Journal of Behavioral Education*, 4, 397-413.
- Cantwell, D. P. y Carlson, G. A.** (1987). *Trastornos afectivos en la Infancia y en la adolescencia*. Barcelona: Martínez Roca.
- Cohen, J. D.** (1966). The Impact of Multivariate Research in Clinical Psychology. En R. B. Cattell (Ed.). *Handbook of Multivariate Experimental Psychology*, Chicago: Rand McNally Co.
- Del Barrio, M. V.** (1984). *Temas de Psicopatología Infantil*. Valencia: Promolibro.
- Elstein, A. S., Shulman, L. S. y Sprafka, S. A.** (1978). *Medical Problem Solving*. Cambridge, Mass: Harvard University Press.
- Felton, J. L. y Nelson, R. O.** (1984). Inter-Assessor Agreement on Hypothesized Controlling Variables and Treatment Proposals. *Behavioral Assessment*, 6, 199-208.
- Hagopian, L.P.; Fisher, W.W.; Thompson, R.H.; Owen, D.J. y Cols.** (1997). Toward the development of structure criteria for interpretation of functional analysis data. *Journal of applied Behavior Analysis*, 30, 313-321.
- Hawkins, R. P.** (1986). Selection of Target Behaviors. En R. O. Nelson y S. C. Hayes (Eds.). *Conceptual Foundations of Behavioral Assessment*. New York: The Guildford Press.
- Hay, W. M., Hay, L. R., Angle, H. V. y Nelson, R. O.** (1979). The Reliability of Problem Identification in the Behavioral Interview. *Behavioral Assessment*, 1, 107-118.
- Haynes, S. N.** (1979). *Principles of Behavioral Assessment*. New York: Gardner Press, Inc.
- Kanfer, F. H. y Saslow, G.** (1969). An Outline for Behavioral Diagnosis. En C. M. Francks (Comp.). *Behavior Therapy: Appraisal and Status*, Nueva York: McGraw-Hill.
- Karoly, P.** (1993). Goal systems: An organizing framework for clinical assessment and treatment planning. Special section: Treatment

- implications of psychological assessment. *Psychological Assessment*, 5, 273-280.
- Lefkowitz, M. M. y Burton, N.** (1978). Childhood Depression: A Critique of the Concept. *Psychological Bulletin*, 85, 716-726.
- Mash, E. J. y Terdal, L. G.** (1979). *Behavioral Assessment of Childhood Disorders*. Nueva York: Guildford Press.
- Messick, S.** (1989). Validity. En R. L. Lynn (Comp.). *Educational Measurement*. Nueva York: American Council of Education y Macmillan Publishing Company.
- Messick, S.** (1993). *Foundations of Validity: Meaning and Consequences in Psychological Assessment*. Conferencia presentada en el "IV Congreso de Evaluación Psicológica" celebrado en Groninga, Agosto de 1993, 1993
- Miller, L. C., Barrett, C. L. y Hampe, E.** (1974). Phobias of Childhood in a Prescientific Era. En A. Davis (Ed.). *Child Personality and Psychopathology: Current Topics*. Nueva York: Wiley.
- Moreno, B. y Valero, L.** (1987). *Casos de Psicología Clínica*. Málaga: Universidad de Málaga.
- Nelson, R. O. y Hayes, S. C.** (1986). The Nature of Behavioral Assessment. En R. O. Nelson y S. C. Hayes (Eds.). *Conceptual Foundations of Behavioral Assessment*. New York: The Guilford Press
- O'Leary, K. D.** (1979). Behavioral Assessment. *Behavioral Assessment*, 1, 31-36.
- O'Leary, K. D. y Wilson, G. T.** (1987). *Behavior Therapy. Application and Outcome*. New Jersey: Prentice-Hall.
- Pelechano, V.** (1981). *Miedos Infantiles y Terapia Familiar Natural*. Valencia: Alfaplus.
- Robins, L.N.** (1979). Follow up studies. En H.C. Quay y J.S. Werry (Eds.). *Psychopathological disorders of childhood*. Nueva York: Wiley.
- Romero, J. F.** (1988). *Trastornos del Aprendizaje I*. Málaga: Universidad de Málaga.
- Romero, J. F., Cano, J., Fuentes, F., Castro, S. y Lopez, D.** (1989). *Trastornos del Aprendizaje III. Dificultades Psicosociales y Emocionales en la Infancia*. Málaga: Universidad de Málaga.
- Shechtman, A.** (1970). Age Patterns in Children's Psychiatric Symptoms. *Child Development*, 41, 683-693.
- Schulte, D.** (1974). Un Esquema para el Diagnóstico y la Planificación Terapéutica en la Modificación de Conducta. En D. Schulte (Dir.). *Diagnostik in der Verhaltenstherapie*, (pp. 75- 104). München: Urban und Schwarzenberg.

- Schulte, D. (1989). Clinical Diagnoses and Treatment Planning. *Evaluación Psicológica/Psychological Assessment*, 5, 135-146.
- Silva, F. (1989). *Evaluación Conductual y Criterios Psicométricos*. Madrid: Pirámide.
- Silva, F. (1993). Treatment Utility: A Reappraisal. *European Journal of Psychological Assessment*, 9, 222-226.
- Suen, H. K.; Ary, D. y Ary, R. M. (1986). A Note on the Relationship Among Eight Indices of Interobserver Agreement. *Behavioral Assessment*, 8, 301-303.
- Swan, G. y Macdonald, M. (1978). Behavior Therapy in Practice: A National Survey of Behavior Therapists. *Behavior Therapy*, 9, 799- 807.
- Turk, D.C. y Salovey, P. (1988). *Reasoning, inference , and judgment in clinical psychology*. Nueva York: Free Press.
- Ward, C. H., Beck, A. T., Mendelson, M., Mock, J. E. y Erbaugh, J. K. (1962). The Psychiatric Nomenclature: Reasons for Diagnostic Disagreement. *Archives of General Psychiatry*, 7, 198-205.
- Weist, M.D.; Ollendick, T.H. y Finney, J.W. (1991). Toward the empirical validation of treatment targets in children. *Clinical Psychology Review*, 11, 515-538.
- Westmeyer, H. y Hageböck, J. (1992). Computer-Assisted Assessment: A Normative Perspective. *European Journal of Psychological Assessment*, 8, 1-16.
- Wilson, F. E. y Evans, E. A. (1983). The Reliability of Target- Behavior Selection in Behavior Assessment. *Behavioral Assessment*, 5, 15-32.