

VALIDEZ A POSTERIORI DE LOS DIAGNOSTICOS CLINICOS

Antonio Godoy

Facultad de Psicología. Universidad de Málaga

La creencia de que cuando se obtienen resultados de éxito tras la aplicación de un tratamiento psicológico se validan tanto el tratamiento aplicado como la evaluación pretratamiento parece encontrarse bien asentada en el campo de la terapia de conducta y la evaluación conductual. Esta creencia, sin embargo, sólo es mantenible cuando se posee evidencia de la validez interna de la intervención psicológica realizada, cosa que muy rara vez ocurre. Debido a ello, la idea de que los resultados positivos validan el diagnóstico y el tratamiento la mayoría de las veces carece de fundamento y puede resultar contraproducente, ya que evita que el clínico aprenda de forma más objetiva acerca de su propia experiencia profesional. Los mecanismos que influyen en que se genere y se mantenga este tipo de creencias son múltiples, destacando entre ellos los siguientes: expresión de las hipótesis diagnósticas de forma acientífica, la tendencia a razonar de forma monocausal y a verificar, más que a refutar, dichas hipótesis, el fenómeno de regresión a la media y la estructura de la situación de aprendizaje que se da en las situaciones clínicas.

Palabras clave: Utilidad del tratamiento. Corrección de la evaluación pretratamiento. Corrección de las inferencias clínicas.

A posteriori validity of clinical diagnosis. The tenet holding that success outcomes from treatment validate applied treatment as well pretreatment assessment is a commonplace in behavior therapy and behavioral assessment. Nevertheless, this belief must be maintained only if there is evidence that any other factor has not produced such success outcomes. Since in psychotherapy other factors frequently can be brought to account for outcomes, this belief is unjustified and sometimes could prevent therapist to learn objectively from experience. Several mechanisms are responsible for maintaining this tenet, mainly unscientific clinical hypotheses, monocausal reasoning, hypothesis substantiation strategies, rather than hypothesis refutation, regression to mean, and the learning structure of therapy setting.

Key words: Treatment utility. Correctness of pretreatment assessment. Correctness of clinical inferences.

EL ÉXITO DE LOS RESULTADOS TRAS LA INTERVENCIÓN CLÍNICA NO IMPLICA LA CORRECCIÓN DEL DIAGNÓSTICO PREVIO.

Ha llegado a ser un tópico o lugar común en evaluación conductual y en modificación de conducta que la mejor evaluación del

diagnóstico se da con el tratamiento (Hayes, Nelson, & Jarret, 1986, 1987; Nelson & Hayes, 1979a, 1979b, 1981, 1986). Esto es, cuando tras la aplicación del tratamiento se obtiene éxito, dicho resultado valida tanto el diagnóstico emitido como el tratamiento aplicado. Así, si el tratamiento tiene éxito, se puede concluir que la evaluación ha sido acertada, tanto en el sentido de que ha sido fiable y válida (en un grado no despreciable) como útil. No se suele decir nada para los casos en los que el tratamiento no tiene éxito, ya que en estas ocasiones parece pensarse

Correspondencia: Antonio Godoy.
Departamento de Psicología Social y de la Personalidad.
Facultad de Psicología.
Universidad de Málaga. 29071. SPAIN

que la situación es inconclusiva, puesto que podría ocurrir que la evaluación fuera incorrecta, que el tratamiento fuera inadecuado, o que ambos, evaluación y tratamiento, fueran desacertados.

Sin embargo, los resultados del tratamiento no pueden utilizarse, sin más, para validar *a posteriori* la evaluación o diagnóstico, ni cuando dichos resultados son positivos ni cuando son negativos.

La creencia de que los resultados positivos validan tanto el diagnóstico como el tratamiento aplicado sólo es defendible si se cree que dichos resultados se han producido a causa de que se ha realizado el diagnóstico y se ha aplicado el tratamiento. Sin embargo, si los resultados que se producen tras la intervención psicológica pueden atribuirse a cualquier otra causa distinta de la intervención del psicólogo, entonces dichos resultados no validan ni el diagnóstico ni el tratamiento. Como estas posibles causas alternativas siempre existen, puede concluirse entonces que, tanto si los resultados del tratamiento son positivos como si son negativos, nunca sirven para validar *a posteriori*, de forma correcta, la calidad y adecuación de la evaluación pretratamiento, y mucho menos la creencia de que los resultados han sido positivos porque el diagnóstico ha sido acertado y el tratamiento efectivo para producir dichos resultados.

Aunque algunos autores han criticado esta forma de validar *a posteriori* los resultados del diagnóstico (Silva, 1978, 1985, 1989; Vizcarro, 1987), este pensamiento parece seguir siendo frecuente en la evaluación psicológica clínica. Debido a ello, vamos a desarrollar a continuación con cierto detenimiento algunas posibilidades que pueden presentarse y las conclusiones lógicas a que es posible llegar, tanto cuando se obtienen resultados positivos como cuando se obtienen resultados negativos tras la aplicación del tratamiento. Aunque no nos vamos a detener aquí en analizar qué pueden significar las expresiones «obtener resultados po-

sitivos o de éxito» y «obtener resultados negativos o de fracaso», no debiera olvidarse, sin embargo, que los resultados pueden ser positivos o negativos tanto en relación con el diagnóstico previamente realizado como en relación a las quejas y demandas presentadas por el sujeto y por su entorno (Godoy & Silva, 1990), así como que los criterios para juzgar el éxito de un determinado tratamiento terapéutico pueden ser múltiples, tanto a nivel substantivo (desaparición de los síntomas, desaparición de los factores que supuestamente producían los síntomas, mejor calidad de vida del cliente o de las personas que lo rodean, un mejor ajuste del yo ideal y el yo real, una mayor autorrealización, aumentar la fuerza del Yo a expensas de la del Super-Yo, etc.) como a nivel metodológico (significación estadística del cambio experimentado, mejoría clínica importante, o relevancia social de los cambios acaecidos en el comportamiento del sujeto). Lo único que se necesita para nuestra exposición es mantener que existe algún método por el que es posible concluir que un determinado resultado es positivo (éxito) o negativo (fracaso). Esta creencia parece ser mantenida por la mayoría de los psicólogos clínicos, que suelen situar los resultados de sus intervenciones en algún punto en el continuo éxito-fracaso (seguramente más cercano al polo éxito que al polo fracaso).

Así pues, cuando un clínico predice que su cliente no se curará si no es con terapia, si luego ocurre que el sujeto se cura tras aplicarle la terapia indicada, la predicción se da por confirmada. Pero también parece ser cierto en la práctica profesional que, si el sujeto no se cura, a pesar del tratamiento, con frecuencia también se termina por pensar que se estaba en lo cierto, quizá no completamente, pero desde luego no del todo equivocado. En algunos casos incluso parece concluirse que aún habría sido peor si no se llega a aplicar el tratamiento. Ocurre lo que ocurra, con frecuencia las predicciones clínicas parecen autoconfirmarse. Algo semejante

sucede con las medidas preventivas de una situación predicha (Einhorn & Hogarth, 1978).

Supongamos que se diagnostica depresión a un individuo que acude a consulta, y que se considera que el factor que mantiene dicha depresión son las ideas irracionales que mantiene acerca de sí mismo (v.gr., «yo no sirvo para nada») y del mundo que lo rodea (v.gr., «todos los demás son superiores a mí»). Sobre la base de este diagnóstico, la predicción que se realiza es, pues, que, si se cambian dichas ideas, la depresión deberá desaparecer. En caso contrario, la depresión seguirá su curso. Ante esta conceptualización del caso, el psicólogo decide aplicar Terapia Racional Emotiva para combatir las ideas irracionales, y además manda tareas en las que el sujeto debe comparar sus capacidades con las de otras personas, de tal forma que compruebe sobre la realidad que no es ni superior ni inferior que muchos otros. Dado el diagnóstico, se espera que, con el tratamiento indicado, dicho paciente mejore su forma de pensar y, como consecuencia, los comportamientos deprimidos por los que ha acudido al psicólogo. Ante estas predicciones pueden ocurrir varios resultados:

Los resultados obtenidos coinciden con los resultados predichos.

Si ocurre lo que se había predicho (si la depresión mejora) se suele concluir que el diagnóstico y el tratamiento, así como la aplicación de éste último, han sido los adecuados. Esto es, el diagnóstico ha sido acertado, y el tratamiento efectivo.

Sin embargo, también es posible pensar en otras alternativas. Antes de exponerlas, no obstante, realizaremos algunas precisiones terminológicas. Aunque en este contexto es usual utilizar de forma intercambiable los términos «correcto», «acertado», «adecuado», y, en el caso del tratamiento «efectivo», en lo que sigue no se entenderán como sinónimos.

Así, «efectivo» se aplicará únicamente al

tratamiento. Se dirá que un tratamiento es efectivo si tras su aplicación se producen resultados positivos o de éxito en el comportamiento del paciente. En caso contrario diremos que el tratamiento ha sido inefectivo. La aplicación de un tratamiento será «adecuada» cuando se realice de acuerdo con las normas que sobre el mismo puedan existir. En caso contrario, la aplicación será «inadecuada». Un tratamiento es «correcto» si se ha derivado del diagnóstico de forma psicológica y lógicamente apropiada, esto es, cuando, dadas las premisas afirmadas en el diagnóstico y dados los conocimientos actuales en psicología, cabe predecir que dicho tratamiento será el más efectivo. Cabe pensar, por tanto, que en algunas ocasiones un tratamiento correctamente derivado del diagnóstico podría resultar, en términos absolutos, inefectivo, bien porque el trastorno presente es bastante resistente a cualquier tipo de tratamiento, bien porque otras variables, ajenas a la terapia, han influido de forma negativa, bien porque el tratamiento se ha aplicado de forma inadecuada. En este sentido, pues, puede decirse que no todos los tratamientos correctamente derivados del diagnóstico y de los conocimientos acumulados en psicología serán efectivos.

A su vez, diremos que un diagnóstico es «acertado» cuando establece exactamente, ni más ni menos, lo que cabe establecer por algún otro tipo de medios no sesgados (o al menos, menos sesgados). Cuando se considera que el diagnóstico emitido contiene errores, bien porque le falta, bien porque le sobra algo, diremos que es «desacertado». El acierto o desacierto de un diagnóstico, entonces, tiene que ver con los conceptos clásicos de fiabilidad y validez. Un diagnóstico es «irrelevante» si, con independencia de su acierto o desacierto, no sirve para establecer un tratamiento, o no influye en el establecimiento del mismo. Diremos, pues, que un diagnóstico es irrelevante si carece de utilidad para establecer el tratamiento subsiguiente (Nelson & Hayes, 1981; Silva,

1989). Aunque es difícil imaginar que un diagnóstico completamente desacertado pueda ser relevante, es fácil pensar en situaciones en las que un diagnóstico acertado (fiable y válido) carece de utilidad práctica en alguna situación particular (bien porque aún no existe tratamiento eficaz alguno, porque los tratamientos existentes no son aplicables en el caso presente, o por otras razones).

Con estas definiciones en mente, pues, veamos algunas posibilidades lógicas que cabe pensar cuando los resultados, tras la realización del diagnóstico y la aplicación del tratamiento, resultan ser positivos o de éxito.

a) El diagnóstico ha sido acertado, el tratamiento correcto y su aplicación adecuada. El tratamiento ha resultado ser efectivo. Sin embargo, que el diagnóstico no sea desacertado y que, al final, se haya producido éxito no quiere decir que se haya actuado de la mejor forma posible. Que los efectos del tratamiento resulten ser positivos no impide pensar que, si se hubieran elegido por ejemplo otras variables controladoras u otras técnicas de tratamiento, hubieran sido aún más positivos, se habrían logrado en menos tiempo, o habrían sido más duraderos o más generalizados.

b) El diagnóstico ha sido el acertado, el tratamiento correcto y su aplicación adecuada, sin embargo, los efectos observados no pueden atribuirse (al menos de forma completa) al tratamiento aplicado. Esto es posible, ya que la corrección del tratamiento indicado y la adecuación de su aplicación no son las únicas variables que influyen sobre la efectividad. Dado que estamos suponiendo que los resultados, al final de la intervención, han sido de éxito, habría que atribuir dicho éxito a variables ajenas a la intervención clínica. Resumidamente, lo que se quiere dar a entender aquí es que los buenos resultados tras la aplicación del tratamiento no por fuerza se deben a lo realizado por el clínico. A esto se refiere la

denominada validez interna de los experimentos y de cualquier intervención o manipulación (Campbell & Stanley, 1963). Que tras la intervención se obtenga un resultado de éxito no justifica la creencia de que dicho resultado se ha producido justo por lo que se ha hecho durante la intervención. Para que esta creencia se convierta en plausible o creíble es necesario además que no existan otras hipótesis alternativas que razonablemente también podrían explicar los resultados.

Así pues, puede mantenerse que el diagnóstico ha sido acertado y que, aunque el tratamiento aplicado ha sido inefectivo, un tercer factor, ajeno a los considerados, ha producido los resultados. Esto es, el sujeto efectivamente sufría depresión y mantenía las ideas irracionales postuladas. Sin embargo, éstas han cambiado no por el tratamiento aplicado, sino por otro tipo de experiencias (v.gr., le ha tocado la lotería; de forma imprevista lo han ascendido en el trabajo, etc.) que le ha llevado a cambiar las ideas y sentimientos previos.

c) El diagnóstico ha sido irrelevante. Se había diagnosticado, de forma desacertada, que el sujeto padecía depresión y se había predicho que, de no aplicar el tratamiento, ésta iba a seguir su curso. Cuando se vuelve a evaluar al sujeto tras aplicar el tratamiento -que puede haber influido o no, en algún sentido, en el sujeto-, se encuentra, de forma acertada, que no padece depresión. Se concluye entonces que el diagnóstico ha sido acertado y el tratamiento adecuado y efectivo. Sin embargo, los resultados de la segunda evaluación, con y sin diagnóstico previo, con y sin tratamiento, habrían sido los mismos. Ni el diagnóstico ni el tratamiento han jugado papel alguno en los resultados obtenidos al final de la terapia. Tanto el diagnóstico como el tratamiento han sido irrelevantes en lo que a la depresión se refiere. Aunque así expuesta, no parece muy razonable esta forma de actuar, con frecuencia se suele establecer la valoración de la inter-

vención comprobando que el diagnóstico previamente asignado al sujeto ya no le resulta aplicable, que ya no presenta los síntomas o comportamientos alterados previos. En estos casos, sin embargo, el cambio en el diagnóstico puede obedecer más a la poca fiabilidad del diagnóstico o del diagnosticador que a los cambios de comportamiento manifestados por el sujeto (Blum, 1978; Mischel, 1968).

d) También podría ocurrir que el diagnóstico ha sido desacertado, pero que el tratamiento, incorrectamente derivado del diagnóstico, ha resultado ser efectivo. Este sería un caso, poco frecuente, en que por azar «se toca la flauta», un caso de equivocación doble con resultado de éxito. Esto puede ocurrir debido, por ejemplo, a que el tratamiento aplicado posee un amplio espectro, de tal forma que es aplicable a un amplio rango de alteraciones, incluida la que de hecho padece el sujeto.

Los resultados obtenidos difieren de los resultados predichos.

Si no ocurre lo que se había predicho (la depresión no mejora), se suele concluir, bien que el diagnóstico ha sido desacertado, bien que las acciones emprendidas han sido inefectivas, bien ambas cosas. Veamos esto con más detenimiento.

a) El diagnóstico ha sido desacertado. El tratamiento por tanto, si se había diseñado de acuerdo con el diagnóstico, probablemente también lo ha sido. En estos casos suele tender a pensarse que, debido a los errores cometidos, toda la intervención clínica, diagnóstico y tratamiento, ha resultado infructuosa. Sin embargo, resulta igualmente posible pensar que tanto el diagnóstico (equivocado) como el tratamiento (contraindicado) no han sido irrelevantes sino que, más bien, han producido un resultado indeseado.

Supongamos, por ejemplo, que el sujeto no sufría depresión, pero que al ponerlo a competir y compararse con otros, sin em-

bargo, ha comenzado a experimentar fracaso tras fracaso, ha generado ideas irracionales que antes no existían y ha terminado por deprimirse. Con la aplicación del tratamiento hemos precipitado aquello de lo que se quería escapar. Esto es, nos hemos equivocado en el diagnóstico y el tratamiento: justo se ha aplicado el tratamiento que producía lo que se quería evitar. Si no se hubiera predicho ni pensado en tratamiento alguno, quizá no hubiera ocurrido el resultado. Este es pues, un caso de equivocación doble (con resultado de fracaso).

b) El diagnóstico ha sido el acertado. Sin embargo, el tratamiento no ha sido el correcto, ha estado mal diseñado, o ha sido mal aplicado.

Lo reflejado en los dos puntos anteriores (a y b) es lo que usualmente tiende a pensarse. No obstante, también es posible pensar en otra eventualidad.

c) El diagnóstico ha sido correcto y el tratamiento el adecuado; sin embargo el resultado ha sido negativo. Si no se ha producido la mejoría o cura prevista no ha sido, por tanto, por errores del clínico al diagnosticar y aplicar el tratamiento. Pensar que si el resultado ha sido de fracaso, por fuerza el diagnóstico o el tratamiento (o ambos) han sido incorrectos supone mantener varias creencias aparejadas que seguramente pocos psicólogos estarían dispuestos a defender en firme: primero, que todos y cada uno de los datos utilizados en la evaluación eran perfectamente fiables y válidos; y segundo, que sólo los tratamientos completamente efectivos en el cien por cien de los casos pueden catalogarse como «adecuados». En otras palabras, el clínico puede haber realizado el mejor diagnóstico posible y haber aplicado de forma impecable el mejor tratamiento disponible y, no obstante, el cliente puede no haber mejorado en absoluto, ya que «el mejor diagnóstico posible» no es cien por cien infalible, ni «el tratamiento disponible más efectivo» cien por cien efectivo. Seguramente todos los psicólogos admiten que,

aunque pueden darse diagnósticos acertados, ningún diagnóstico es infalible. De la misma forma, aunque contamos con técnicas de tratamiento bastante poderosas, ninguna de ellas es completamente efectiva en todos los casos.

Aparte de lo anteriormente manifestado, también es posible pensar que el diagnóstico ha sido acertado y que, además, el tratamiento, dado el diagnóstico realizado, era el adecuado. Sin embargo, otras causas ajenas a la terapia (v.gr., la muerte de un ser querido; un fracaso laboral o familiar grave, etc.) han impedido que el sujeto salga de la depresión. También en estos casos el diagnóstico y el tratamiento habrían sido válidos y, no obstante, los resultados obtenidos habrían resultado negativos.

Resumiendo lo dicho hasta ahora, pues, cabe concluir que, ni cuando los resultados de la intervención psicológica son positivos, ni cuando son negativos pueden validarse con seguridad el diagnóstico y el tratamiento. En el primer caso porque, cuando el resultado es positivo, el diagnóstico puede ser acertado (como en el anterior punto 1.a), desacertado (como en el punto 1.d), o irrelevante (punto 1.c). De la misma forma, el tratamiento puede ser efectivo (punto 1.a), o inefectivo (punto 1.b). En el segundo caso, cuando el resultado es negativo, el diagnóstico puede ser igualmente acertado (punto 2.b) o desacertado (punto 2.a), y el tratamiento efectivo (punto 2.a) o inefectivo (punto 2.c).

¿POR QUÉ SE MANTIENE LA CREENCIA DE QUE EL ÉXITO IMPLICA UN DIAGNÓSTICO ACERTADO?

Posiblemente influyen muchas variables en que se mantenga de forma tan generalizada la creencia de que, si se obtiene éxito, la intervención realizada ha quedado validada. En lo que sigue se expondrán sólo algunas variables y razonamientos que nos han parecido especialmente pertinentes.

Creencia acerca de que el diagnóstico es útil para derivar el tratamiento indicado.

A pesar de las múltiples llamadas de atención acerca de la falta de relación entre diagnóstico y tratamiento (Barrios, 1988; Hartman, Roper, & Bradford, 1979; McReynolds, 1985; Meehl, 1973), la idea de que la evaluación conductual es de gran utilidad para la posterior modificación del comportamiento del sujeto parece encontrarse bien asentada, hasta el punto de que muchos percibirían la aplicación del tratamiento como un sinsentido si previamente no se ha realizado una evaluación previa pormenorizada. De esta forma, se da por supuesto que el tratamiento indicado y aplicado siempre es una derivación lógico-psicológica de la evaluación pretratamiento, de tal forma que al validar el tratamiento (desde la consideración de los resultados terapéuticos obtenidos) también se valida la evaluación de la que dicho tratamiento se ha derivado. Es pues esta idea de que el diagnóstico y el tratamiento se encuentran estrechamente ligados (y no sólo en sentido temporal) uno de los factores que han contribuido a generar y mantener la idea de que los resultados positivos de la intervención psicológica validan no sólo el tratamiento que supuestamente los ha producido, sino también la evaluación de la que el tratamiento aparentemente se ha derivado.

Desde el estado actual de nuestros conocimientos en psicología pueden aducirse, además, otras muchas razones para explicar por qué es tan general la creencia de que los resultados de éxito tras las intervenciones clínicas (o de otro tipo) implican que se ha realizado un diagnóstico acertado y un tratamiento correcto y efectivo. Veamos algunas de ellas.

Las hipótesis diagnósticas son tan imprecisas que no hay información que pueda refutarlas.

En primer lugar, es posible que el clínico esté actuando de forma acientífica. Esto es,

que esté manteniendo hipótesis diagnósticas no falsables, en el sentido de K. Popper (1982), hipótesis que son compatibles con cualquier comportamiento que pueda presentar el paciente. También es posible que las hipótesis diagnósticas que formula sí sean falsables en principio, pero que con frecuencia se las inmunice contra una posibilidad efectiva de falsación, ya que si ocurre lo que se predice, la hipótesis se confirma, y si no ocurre, se modifica con razonamientos *ad hoc* e introducción de terceras variables antes no aludidas (otras influencias externas, dificultades en la aplicación del tratamiento, falta de motivación del sujeto, negativa a colaborar, etc.).

Queremos pensar, no obstante, que la situación más usual es aquélla en la que se trabaja manteniendo en mente el criterio de falsabilidad que, aunque muy criticado por otros filósofos de la ciencia, parece mantener un cierto predicamento entre los psicólogos. Sin embargo, aún suponiendo que se trabaja con hipótesis científicas, falsables, y que se está dispuesto a aceptar que son falsas, hay varios fenómenos cognitivos y metodológicos que hacen poco probable que se aprenda de la experiencia profesional cotidiana algo distinto que la idea de que el éxito de la intervención valida el diagnóstico y el tratamiento realizados. Sin ánimo de ser exhaustivos, veamos algunos de estos fenómenos.

Tendencia a razonar de forma monocausal.

Entre los fenómenos cognitivos a que antes hemos hecho referencia cabe destacar sin duda la tendencia de las personas humanas (al menos en la sociedad occidental) a razonar en términos de relaciones causales, un resultado que ha aparecido con frecuencia tanto en las investigaciones sobre razonamiento diagnóstico como en otras muchas situaciones (Ajzen, 1977; Einhorn, 1986). Muchos clínicos parecen actuar como si pensarán que las causas implican sus efectos,

pero también a la inversa, que los efectos implican sus causas, de tal forma que conociendo el efecto es posible remontarse a las causas que lo han producido. Esto es, a efectos formales puede decirse que tiende a razonarse sobre el condicional como si se tratara de un bicondicional (Frías & Godoy, 1993; Wason & Johnson-Laird, 1972). Con frecuencia parece olvidarse que, en psicología al menos, existen multitud de causas suficientes, pero no necesarias, y que un resultado puede ser el efecto de múltiples causas.

Tendencia a verificar las hipótesis y a recoger información redundante.

Otro fenómeno cognitivo frecuentemente encontrado en las personas en general, y en el razonamiento clínico en particular, es una fuerte tendencia a verificar las propias hipótesis, no a falsarlas (Elstein, Shulman, & Sprafka, 1978; Frías & Godoy, 1993; Mynatt et al., 1977, 1978; Wason & Johnson-Laird, 1972). En general puede decirse que las personas tendemos, no a someter a prueba nuestras propias hipótesis, sino más bien a lograr seguridad en el propio juicio acumulando evidencia a favor del mismo (Baron, Beattie, & Hershey, 1988), lo que tiende a producir una gran cantidad de información altamente redundante y coincidente con la hipótesis formulada (Einhorn, 1986). Ante tal cantidad de información a favor de la hipótesis propuesta, el clínico tiende a ganar seguridad en su propio juicio (Arkes, 1991; Goldberg, 1968; Kahneman & Tversky, 1973; Mischel, 1968, pp. 127-129; Nisbett & Ross, 1980, pp. 119-120; Slovic, Fischhoff, & Lichtenstein, 1977), olvidando que la información redundante no mejora su opinión diagnóstica, y tanto menos cuanto más redundante resulte la información con que cuenta, ya que, como es sabido, cuanto más alta es la correlación entre los predictores tanto menos se gana en la predicción del criterio.

Por su parte, la información discordante,

al contrastar con una gran cantidad de datos redundantes que señalan en la misma dirección, tiende a minusvalorarse o a despreciarse, posiblemente porque se la considera poco fiable (Anderson & Jacobson, 1965; Friedlander & Phillips, 1984; Hoffman, 1968; Koriat, Lichtenstein, & Fischhoff, 1980; Slovic, 1966; Slovic, Fischhoff, & Lichtenstein, 1977).

El fenómeno de regresión a la media.

Dado que el fenómeno de regresión a la media parece ser de difícil comprensión para muchas personas, nos detendremos brevemente en exponer su naturaleza. En la Figura puede observarse una representación del fenómeno de regresión a la media cuando se pretende predecir la respuesta a un determinado tratamiento partiendo de un diagnóstico concreto, supongamos, por seguir con el ejemplo previo, de depresión. El sujeto ha recibido el diagnóstico de depresión, entre otras consideraciones, porque, como puede apreciarse, la puntuación pretratamiento del sujeto en depresión es elevada (D), por encima de la media de la población (N). Las letras X, Y y Z representan tres resultados posibles, tres valores en la medición postratamiento de la depresión. El óvalo representa que la relación entre el diagnóstico y los resultados (la relación entre la evaluación pre- y la postratamiento de la depresión) no es perfecta, bien porque el diagnóstico y la evaluación de dichos resultados no es completamente fiable, bien porque no siempre se aplica de forma idéntica la técnica de tratamiento o no siempre se aplica el mismo tratamiento, bien porque la relación entre el tratamiento y los resultados obtenidos es imperfecta (por la acción, por ejemplo, de terceras variables que también influyen sobre los resultados). La línea continua representa la relación perfecta entre el diagnóstico y los resultados de la intervención, relación perfecta que sólo se producirá si los resultados dependen completamente del diagnóstico realizado y de ninguna

otra variable (por ejemplo, falta de motivación del sujeto, inadecuada aplicación de las técnicas de tratamiento, remisión espontánea, etc.). El óvalo se reduciría a dicha línea si la relación entre el diagnóstico y los resultados fuera perfecta. Esto es, cuando sucesivas evaluaciones producen exactamente el mismo valor diagnóstico, siempre se aplica la misma técnica de tratamiento exactamente de la misma forma, el tratamiento aplicado es cien por cien efectivo, no existen otras variables que puedan favorecer o entorpecer los efectos del tratamiento, etc. Dado que toda esta serie de condiciones no suelen darse nunca juntas, en la práctica nunca contaremos con una relación perfecta, sino más bien con la relación representada mediante el óvalo. Cuando la relación es perfecta, a un sujeto que obtenga el diagnóstico D se le predecirá exactamente el valor Y de los resultados. Sin embargo, cuando dicha relación no es perfecta, como suele suceder, al sujeto con el diagnóstico D se le puede atribuir cualquier puntuación entre X y Z o, si se prefiere, entre a y b. Por tanto no hay un único valor exacto de los resultados que se le pueda atribuir con precisión, sino cualquier valor entre a y b. En la Figura puede apreciarse también cómo el segmento o-a es mayor que el segmento o-b. Esto refleja el hecho de que existen más valores de los resultados por debajo de o (e Y) que por encima. Aunque el sujeto podría obtener cualquier resultado entre a y b (o lo que es lo mismo, entre X y Z), lo más probable, por tanto, es que obtenga un resultado entre o y a, es decir y en nuestro ejemplo, una depresión postratamiento más baja que la que cabría predecir si la relación fuera perfecta. Este es el fenómeno que se conoce como regresión a la media: la probabilidad de obtener una puntuación postratamiento entre o y a es mayor que la de obtener una puntuación entre o y b, o, lo que es lo mismo, lo más probable es que el sujeto obtenga una puntuación en depresión por debajo de Y, más cercano a la media, aparentando que la de-

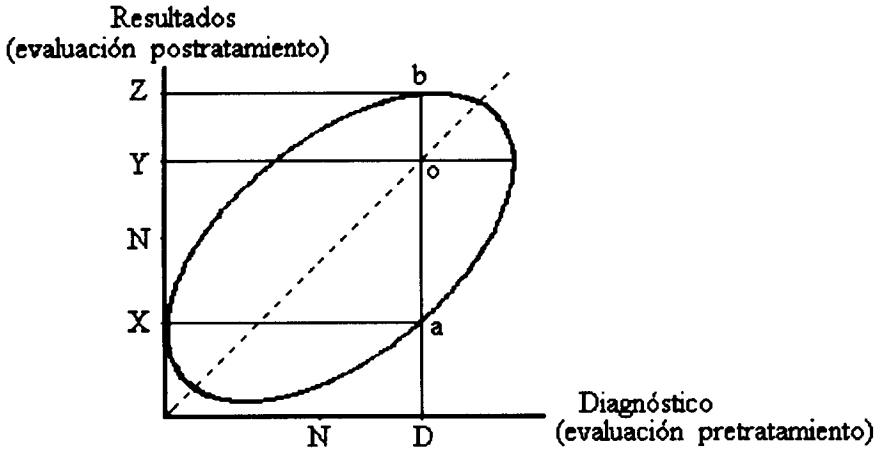


Figura: Efecto de regresión hacia la media.

presión del sujeto es ahora más normal de lo que era en la evaluación pretratamiento.

Por tanto, desde este punto de vista puede decirse que, si el diagnóstico no es totalmente fiable, o si los resultados observados pueden ser producidos, además de por el tratamiento aplicado, por otros tipos de tratamiento y, en general, por terceras variables, lo que disminuye el valor de la relación entre diagnóstico y resultados obtenidos, cabe esperar que se produzca regresión hacia la media en las conductas problemáticas, y, por tanto, hay cierta probabilidad de que la intervención, aún cuando el diagnóstico y el tratamiento sean inadecuados, parezca un éxito, de tal forma que se concluye que evaluación y tratamiento han quedado validados.

Así, cuando la evaluación es infiable, cabe esperar que el sujeto obtenga puntuaciones más elevadas de lo debido en algunas variables, y menos elevadas en otras. Dado que se toman como conductas problemáticas únicamente las más extremas, se atenderá sólo a las más bajas (v.gr., falta de habilidades sociales) y a las más altas (v.gr., depresión), lo que aumenta la probabilidad de que, por regresión a la media, las más bajas aumenten y las más altas disminuyan,

concluyendo que el tratamiento ha producido éxito, ya que han aumentado las habilidades sociales y ha disminuido la depresión. Las conductas en torno a la media, dado que se dan como normales, no se tienen posteriormente en cuenta. Si se las tuviera, se vería que también se mueven, mejorando algunas y empeorando otras. Si bien las que mejoran correrían el peligro de interpretarse como «efectos de generalización del tratamiento», las que empeoran deberían hacer pensar al terapeuta. Así pues, con una evaluación deficiente y un tratamiento inefectivo es posible concluir que ambos han quedado validados.

La ilusión de validez y el clínico que nunca falla.

A la vista de los dos factores señalados (tendencia a reunir información redundante para verificar la hipótesis formulada y utilización de valores extremos), lo más probable es que el clínico crea, de forma ilusoria pero firme, que su diagnóstico y su tratamiento poseen un alto grado de validez (Einhorn, 1986; Hogarth, 1980). De esta forma, la tendencia a verificar las hipótesis diagnósticas y de tratamiento y la propia naturaleza de las conductas sobre las que se

suele intervenir (con valores extremos, altos o bajos) producen que aparezca el fenómeno denominado «ilusión de validez» (Kahneman & Tversky, 1973). Así pues, cuanto peor actúe el clínico (cuanto más redundante e in fiable sea la información que maneja), tanto mejor se percibirá a sí mismo como clínico y tanto más le demostrará la experiencia que actúa de forma válida.

La existencia de técnicas de tratamiento de amplio espectro.

Vemos, pues, cómo con información in fiable puede concluirse que la intervención realizada es correcta. Pero supongamos que el diagnóstico no sólo resulta poco fiable, sino que es completamente desacertado. Volvamos sobre el último ejemplo propuesto. Supongamos que el sujeto no tiene ni falta de habilidades sociales ni depresión, y que el diagnóstico acertado debería haber sido «falta de reforzamiento en el medio», o cualquier otro (v.gr., cogniciones problemáticas, atribuciones inadecuadas, pérdida del objeto amado, una gran diferencia entre su yo ideal y su yo real, etc.). ¿Qué pasaría si se le aplica un tratamiento poderoso de las habilidades sociales y la depresión? Posiblemente el sujeto logre extraer el poco refuerzo existente en el medio, o incluso logre cambiar de medio. En este caso, el diagnóstico ha sido desacertado, sin embargo el tratamiento aplicado ha resuelto los problemas que se le planteaban al sujeto. Si se evalúa la bondad de la intervención entera se encontrará que, aunque el diagnóstico realizado ha sido desacertado, posiblemente tendamos a dar como buenos tanto el diagnóstico como el tratamiento.

Así pues, cuanto más poderosa resulta una técnica de tratamiento (como puede ocurrir con el ensayo de conducta o con las técnicas de afrontamiento), tanto más se corre el peligro de concluir que toda la actuación llevada a cabo, incluyendo la evaluación, ha quedado validada.

De la misma forma, cuanto peor sea la

evaluación, cuanto más in fiable, tanto más probabilidades de que aparezcan «mejorías» aparentes -regresiones hacia la media- que, según han puesto de manifiesto algunos autores (v.gr., Hogarth, 1980), suelen atribuirse al tratamiento. Cuanto más fiable sea la evaluación, tanto menos probable es que, por azar, se den mejorías. Por tanto, de un diagnóstico in fiable cabe esperar éxito en más ocasiones que de uno fiable, a igualdad de efectividad del tratamiento.

La estructura de las situaciones de aprendizaje a través de la experiencia clínica.

Lo anterior nos hace pensar que un terapeuta, sometido a este tiempo de experiencias terapéuticas, a esta situación de aprendizaje, y dando por supuesto que los resultados finales de éxito validan tanto la evaluación como el tratamiento, lo único que puede es aprender creencias incorrectas (Brehmer, 1980). De esta forma, la validez aparente del tratamiento y de la evaluación continuamente se refuerza, dando lugar a una alta confianza o seguridad en la corrección de las propias actuaciones, a la ilusión de validez a que se han referido algunos autores (Kahneman & Tversky, 1973; Kahneman, Slovic, & Tversky, 1982). Para aprender lo que es correcto debería atenderse a los efectos principales del tratamiento y a los de la evaluación y, por tanto, a qué ocurre con los resultados del tratamiento tanto cuando la evaluación es buena como cuando es mala -o no hay evaluación.

Al no contar con grupo control, o con un control adecuado cualquiera, para diferenciar los efectos del tratamiento de los efectos de otras posibles variables que puedan producir también efectos positivos, a la mejoría producida por los efectos del tratamiento y de la regresión hacia la media hay que añadir los efectos de esas otras variables, que muy probablemente se atribuirán al tratamiento (Einhorn & Hogarth, 1978). Sólo en aquellos casos en que esos otros

factores ajenos al tratamiento produzcan la misma cantidad de efectos positivos que de efectos negativos ambos tipos de efectos se anularán. Sin embargo, los denominados efectos inespecíficos de la intervención permanecerán. De esta forma, en los casos con resultados positivos se pueden estar mezclando los efectos específicos producidos por el tratamiento, los producidos por factores del medio natural ajenos a la terapia –remisión espontánea–, los inespecíficos del tratamiento y los producidos por la regresión hacia la media.

Cuando un terapeuta concluye dando un diagnóstico, suponemos que posee una cierta seguridad de que dicho diagnóstico es válido, ya que de no ser así probablemente seguiría investigando. Sin embargo, cuando se juzga que el diagnóstico es válido, es difícil observar los resultados de los diagnósticos alternativos que podrían haberse formulado, así como pensar en los resultados probables de los tratamientos que no se han aplicado. Esta información limitada hace que uno reciba *feedback* de éxito a pesar de, más bien que debido a, la propia habilidad diagnósticadora (Einhorn y Hogarth, 1978, 1981).

Para referirse a este fenómeno, semejante al aprendizaje supersticioso (Skinner, 1948), se ha propuesto el concepto de «estructuras de aprendizaje de resultados irrelevantes» (Einhorn, 1980). Este concepto se refiere a que en determinadas tareas la información positiva que se recibe puede resultar irrelevante, o incluso contraproducente, para corregir los juicios inadecuados cuando falta la estructura de la tarea, o cuando ésta es errónea. Para ilustrarlo, Einhorn y Hogarth (1981) proponen un ejemplo que resumimos a continuación.

Supongamos el caso de un camarero que piensa que la cuantía de las propinas que le dan depende de la apariencia de los clientes y que él (aficionado a la psicología) es capaz de juzgar dicha apariencia y, por tanto, predecir quién dejará buenas y quién malas propinas. Si el servicio que presta a los dis-

tintos clientes es acorde con dicha creencia, posiblemente la experiencia le demuestre que está en lo correcto, ya que los clientes sobre los que ha predicho que dejarán buenas propinas (a los que atiende bien), efectivamente dejan mejores propinas que los que ha predicho que dejarán poca propina (a los que ha atendido mal). La única forma que tiene de aprender lo que realmente ocurre es experimentar, prestando un buen servicio a algunos de los que parece que van a dejar poca propina, y un mal servicio a algunos de los que parece que van a dejar buena propina. Este experimento posiblemente le enseñe si está o no equivocado. Pero si no lo está, como piensa, le costará dinero y esfuerzo: el dinero que perderá al atender mal a los clientes que supuestamente van a dejar buenas propinas, y el esfuerzo probablemente inútil al atender bien a los que muy posiblemente no dejen propina.

Este «dilema del camarero» suele aparecer con frecuencia en la vida diaria: aprender cuesta esfuerzo y, a veces, puede costar también dinero, prestigio, malos resultados de la terapia, abandono del paciente, y otras consecuencias no deseadas. Estos requerimientos exigidos para que se produzca un aprendizaje correcto de la propia efectividad posiblemente hacen poco probable que el clínico pueda aprender correctamente de su práctica profesional diaria, llevándolo a una situación en la que lo único que se puede aprender son ideas incorrectas (aunque quizá no del todo desadaptativas o insanas) acerca de la propia actuación (Brehmer, 1980; Einhorn, 1980, Einhorn & Hogarth, 1981; Hammond, 1978). Como ha manifestado Brehmer (1980), si el clínico ha de formarse una idea realista de su efectividad como profesional, desde luego ésta no va a provenir de su experiencia directa con los resultados de sus terapias.

Lo más probable, pues, es que el terapeuta obtenga resultados positivos, actúe como actúe, e incluso si no aplica tratamiento alguno, tal como muestran los datos de la re-

misión espontánea (Eysenck, 1952), y tanto si evalúa bien como si evalúa mal, pero especialmente en este último caso. Luego, posiblemente concluya que su actuación es efectiva y vaya ganando confianza en sus propios juicios conforme gana en experiencia, sean dichos juicios y actuaciones adecuados o inadecuados.

Olvido de los casos de remisión espontánea.

Por otra parte, ignorar los efectos de la remisión espontánea es olvidar la tasa base de las mejorías. La no utilización de la tasa base es algo bastante común (Lyon & Slovic, 1976; Nisbett, Borgida, Crandall, & Reed, 1976; Tversky & Kahneman, 1974). La influencia de la remisión espontánea, o tasa base de las remisiones, es importante para extraer conclusiones sobre el propio éxito y para establecer conclusiones correctas de confianza en los propios juicios. Si un terapeuta concluye que el 80% de las conductas problemáticas de un paciente han desaparecido tras la terapia, quizá concluya que ésta ha sido altamente efectiva. Si la tasa de remisión espontánea es del 20%, quizá sea una buena conclusión. Pero si la tasa de la remisión espontánea es del 90%, debería concluir que, a pesar de que la mayoría ha mejorado, el tratamiento es contraproducente, ya que disminuye el número de conductas que mejoran en comparación con la situación en la que no se trata o en la que no se acude a terapia.

Con una experiencia sobre datos dependientes de los propios juicios y acciones, y sin poder disponer de los datos que se habrían producido si se hubiera decidido y hecho lo contrario, lo único que se puede aprender es que las propias acciones son la mayoría de las veces correctas y, en consecuencia, ganar confianza en lo que uno decide y hace (Brehmer, 1980; Brunswik, 1943, 1952; Winkler & Murphy, 1973).

Resumiendo, pues, puede decirse que establecer la utilidad de un diagnóstico para el

tratamiento es una tarea difícil, que sólo puede alcanzarse con los diseños experimentales adecuados. La experiencia que suele resultar posible en la práctica profesional cotidiana, sin embargo, no parece ser la más idónea para aprender si los propios diagnósticos (y tratamientos) son o no adecuados.

COMO DEBE COMPROBARSE LA BONDAD DEL DIAGNOSTICO (Y DEL TRATAMIENTO).

Lo anteriormente dicho pretende poner de manifiesto que la valoración de los efectos de la evaluación clínica y del tratamiento es difícil de llevar a cabo, especialmente cuando se quiere diferenciar entre la corrección del diagnóstico y la del tratamiento. Como ha podido comprobarse, resultados positivos pueden provenir tanto de un diagnóstico acertado como equivocado. De igual forma, resultados negativos pueden provenir de un diagnóstico acertado o desacertado.

Por ello, para deslindar la corrección del diagnóstico de la corrección del tratamiento debe acudir a la investigación y, más específicamente, al empleo de los diseños experimentales adecuados, tal como antes hemos indicado. En concreto, debe utilizarse al menos un diseño factorial 2×2 , en el que hay unos sujetos que se diagnostican y otros que no se diagnostican (o en el que unos tienen un diagnóstico acertado y otros un diagnóstico desacertado), en el que unos reciben tratamiento y otros no lo reciben (o reciben tratamiento correcto unos e incorrecto otros). De este modo se pueden calcular los efectos principales del diagnóstico, del tratamiento y de su interacción. Si no se hace así es imposible deslindar los efectos del diagnóstico y del tratamiento. De esta forma, si los efectos del tratamiento fueran potentes (v.gr., si los efectos de mejorar la forma de pensar del sujeto influyen en una gran cantidad de áreas de la vida), uno esperará obtener unos resultados de cierta valía incluso

en los casos en los que el diagnóstico resulta deficiente o incluso completamente desacertado. Esto es, el tratamiento mostrará efectividad tanto si se diagnostica depresión como si se diagnostica ansiedad, fobia, obsesiones o esquizofrenia.

Así pues, para concluir sobre los efectos del tratamiento, necesitamos conocer qué ocurre cuando no se aplica tratamiento (o, lo que es mejor aún, qué ocurre cuando, dado el diagnóstico realizado, se aplica un tratamiento claramente incorrecto o que debiera resultar inefectivo). Del mismo modo, para conocer los efectos de un buen diagnóstico se precisa conocer lo que sucede cuando no se dispone de diagnóstico alguno, o bien cuando se utiliza un diagnóstico deficiente o claramente desacertado (Einhorn & Hogarth, 1978).

Cuando no disponemos de los datos que señalan qué sucede cuando se aplica un diagnóstico desacertado, o qué ocurre cuando se aplica un tratamiento incorrecto, el extraer conclusiones se complica, ya que no podemos saber si lo que de hecho ha ocurrido habría ocurrido de todas formas, con o sin diagnóstico acertado, con o sin tratamiento correcto. Puede concluirse, por tanto, que establecer una relación segura entre los resultados obtenidos y el diagnóstico (o el tratamiento) utilizado es una tarea que usualmente escapa al clínico en su trabajo cotidiano. Para establecer dicha relación por fuerza necesitamos disponer de datos de los que carece el clínico aplicado: qué ocurre cuando ni se diagnostica ni se trata, o qué cuando se diagnostica otro trastorno o se aplica otro tratamiento. Quizá es necesario admitir que el establecimiento de la utilidad de los diagnósticos y los tratamientos no es algo que le corresponde al clínico aplicado, que precisamente por serlo carece de los datos necesarios, sino más bien a la investigación clínica. Es obvio que también en la práctica clínica cabe realizar investigación. Pero en este caso deberemos estar dispuestos a asumir los riesgos del camarero, tal co-

mo se ha expuesto en el ejemplo anterior: deberemos disponer las cosas de tal forma que algunos pacientes reciban buenos diagnósticos y buenos tratamientos; otros, buenos diagnósticos y malos tratamientos; unos terceros, malos diagnósticos y buenos tratamientos; y otros pacientes, por último, malos diagnósticos y malos tratamientos.

A la luz de lo que se ha dicho a lo largo de estas páginas, pues, parece poder afirmarse que la creencia de que los resultados de éxito de una intervención psicológica validan el diagnóstico y el tratamiento previamente realizados no solamente no es acertada sino que, en varios sentidos, es contraproducente, ya que inmuniza contra un aprendizaje más realista acerca de los propios logros clínicos, impidiendo que puedan mejorarse los fallos y errores cometidos. Es posible, no obstante, que una visión optimista de nuestras actuaciones clínicas no sea del todo psicológicamente desadaptativa y que nos proteja de tener que enfrentarnos con resultados no tan buenos como los que nos gustaría admitir.

REFERENCIAS

- Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology*, 35, 303-314.
- Anderson, N. H., & Jacobson, A. (1965). Effect of stimulus inconsistency and discounting instructions in personality impression formation. *Journal of Personality and Social Psychology*, 2, 531-539.
- Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin*, 110, 486-498.
- Baron, J., Beattie, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning. II. Congruence, information, and certainty. *Organizational Behavior and Human Decision Processes*, 42, 88-110.
- Barrios, B. A. (1988). On the changing nature of behavioral assessment. En A. S. Bellack &

- M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (3.^a ed.), Nueva York: Pergamon.
- Bechtold, H. P. (1971). La validez de constructo: una crítica. En E.I. Megargee (Ed.), *Métrica de la personalidad*. Vol. I, (pp. 162- 180). México: Trillas. (Original inglés de 1966).
- Blum, J. D. (1978). On changes in psychiatric diagnosis over time. *American Psychology*, 11, 1017-1031.
- Brehmer, B. (1980). In one word: Not from experience. *Acta Psychologica*, 45, 223-241.
- Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review*, 50, 255-272.
- Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago: University of Chicago Press. Traduc. castellana: El marco conceptual de la psicología. Madrid: Debate, 1989.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally College Publications. Traduc. castellana: Diseños experimentales y cuasi-experimentales para la investigación en la enseñanza. Buenos Aires: Amorrortu, 1970.
- Einhorn, H. J. (1980). Learning from experience and suboptimal rules in decision making. En T.S. Wallsten (Ed.), *Cognitive processes in choice and decision behavior*, (pp. 66-80). Hillsdale, NJ: Lawrence Erlbaum.
- Einhorn, H. J. (1986). Accepting error to make less error. *Journal of Personality Assessment*, 50, 387-395.
- Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review*, 85, 395-416.
- Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology*, 32, 53-88.
- Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medical Problem Solving. An analysis of clinical reasoning*. Cambridge, Mass.: Harvard University Press.
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology*, 16, 319-324.
- Frías, A., & Godoy, A. (1993). Applying modal logic to diagnostic reasoning. *Second European Conference on Psychological Assessment*. Groningen (Holanda).
- Friedlander, M. L., & Phillips, S. D. (1984). Preventing anchoring errors in clinical judgment. *Journal of Consulting and Clinical Psychology*, 52, 366-371.
- Godoy, A., & Silva, F. (1990). El proceso diagnóstico como una actividad técnico-científica. *II congreso del Colegio Oficial de Psicólogos* Vol. 7: Diagnóstico y Evaluación Psicológica, (pp. 65- 69). Valencia: Colegio Oficial de Psicólogos.
- Goldberg, L. R. (1968). Simple models or simple processes? Some research in clinical judgment. *American Psychologist*, 23, 486- 496.
- Hammond, K. R. (1978). Toward increasing competence of thought in public policy formation. En K.R. Hammond (Ed.), *Judgment and decision in public policy formation*, (pp. 11-32). Denver: Westview.
- Hartman, D. P., Roper, B. L., & Bradford, D. C. (1979). Some relationships between behavioral and traditional assessment. *Journal of Behavioral Assessment*, 1, 3-21.
- Hayes, S. C., Nelson, R. O., & Jarret, R. B. (1986). Evaluating the quality of behavioral assessment. En R.O. Nelson, & S.C. Hayes (Eds.), *Conceptual foundations of behavioral assessment*, (Cap. 13). Nueva York: The Guilford Press.
- Hayes, S. C., Nelson, R. O., & Jarret, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist*, 42, 963-974.
- Hoffman, P. J. (1968). Cue-consistency and configurality in human judgment. En B. Kleinmuntz (Ed.), *Formal representation of human judgment*. Nueva York: John Wiley.
- Hogarth, R. M. (1980). *Judgment and choice. The psychology of decision*. Nueva York: John Wiley.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 251-273.
- Koriat, A., Lichtenstein, S., & Fischhoff, B.

- (1980). Reasons for confidence. *Journal of Experimental Psychology (Human Learning and Memory)*, 6, 107-118.
- Lyon, D., & Slovic, P. (1976). Dominance of accuracy information and neglect of base rates in probability estimation. *Acta Psychologica*, 40, 287-298.
- McReynolds, P. (1985). Psychological assessment and clinical practice: Problems and prospects. En J.N. Butcher, & C.D. Spielberger (Eds.), *Advances in personality assessment* Vol. Vol. 4, Hillsdale, NJ: LEA.
- Meehl, P. E. (1973). Some ruminations on the validation of clinical procedures. En P. E. Meehl (Ed.), *Psychodiagnostic: Selected papers*, Londres: Oxford University Press.
- Mischel, W. (1968). *Personality and assessment*. Nueva York: John Wiley. Traduc. castellana: Personalidad y evaluación. México: Trillas, 1973.
- Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inference. *Quarterly Journal of Experimental Psychology*, 29, 85-95.
- Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1978). Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology*, 30, 395-406.
- Nelson, R. O., & Hayes, S. C. (1979a). Some current dimensions of behavioral assessment. *Behavioral Assessment*, 1, 1-16.
- Nelson, R. O., & Hayes, S. C. (1979b). The nature of behavioral assessment: A commentary. *Journal of Applied Behavior Analysis*, 12, 491-500.
- Nelson, R. O., & Hayes, S. C. (1981). Nature of behavioral assessment. En M. Hersen, & A.S. Bellack (Dir.), *Behavioral assessment: A practical handbook* (2), Nueva York: Pergamon.
- Nelson, R. O., & Hayes, S. C. (1986). The nature of behavioral assessment. En R.O. Nelson, & S.C. Hayes (Eds.), *Conceptual foundations of behavioral assessment*, (Cap. 1). Nueva York: The Guilford Press.
- Nisbett, R. E., Borgida, E., Crandall, R., & Reed, H. (1976). Popular induction: Information is not necessarily informative. En J. S. Carroll, & J.W. Payne (Eds.), *Cognition and social behavior*, (pp. 113-133). Hillsdale, NJ: Lawrence Erlbaum.
- Nisbett, R. F., & Ross, L. (1980). *Human inference: Strategies and shortcoming of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Popper, K. (1982). *La lógica de la investigación científica*. Madrid: Tecnos. (Original alemán de 1934).
- Silva, F. (1978). El análisis funcional de conducta como disciplina diagnóstica. *Análisis y Modificación de Conducta*, 4, 28-55.
- Silva, F. (1985). *Psicodiagnóstico. Teoría y aplicación*. Valencia: Centro Editorial de Servicios y Publicaciones Universitarias.
- Silva, F. (1989). *Evaluación conductual y criterios psicométricos*. Madrid: Pirámide.
- Skinner, B. F. (1948). «Superstition» in the pigeon. *Journal of Experimental Psychology*, 38, 168-172.
- Slovic, P. (1966). Cue consistency and cue utilization in judgment. *American Journal of Psychology*, 79, 427-434.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology*, 28, 1-39.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristic and biases. *Science*, 185, 1124-1131. Traduc. castellana en M. Carretero y J.A. García Madruga (Comps.): *Lecturas de psicología del pensamiento*. Madrid: Alianza, 1984 (Cap. 9).
- Vizcarro, C. (1987). Aproximaciones empíricas al estudio del proceso diagnóstico. *Evaluación Psicológica/Psychological Assessment*, 3, 299-344.
- Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning*. Londres: B.T. Batsford Ltd. Traduc. castellana en Madrid: Ed. Debate, 1981.
- Winkler, R. L., & Murphy, A. H. (1973). Experiments in the laboratory and the real world. *Organizational Behavior and Human Performance*, 10, 252-270.

Aceptado el 1 del 3 de 1994