

Validez de las teorías implícitas de personalidad de profesores de Educación General Básica

JOSÉ LUIS ALVAREZ CASTILLO
Universidad Pontificia de Salamanca



Resumen

Cuatro muestras análogas de profesores ayudaron a comprobar las hipótesis sobre la validez de sus teorías implícitas de personalidad (TIP). El criterio de validez estuvo basado en juicios emitidos inmediatamente después de observar varias secuencias de conductas grabadas en vídeo. En un nivel agregado de análisis, las TIP demostraron poseer fidelidad estructural al correlacionar con las covariaciones reales entre las categorías, al mismo tiempo que contribuyeron a la validez de los juicios sociales. Por otra parte, la semántica no demostró ser un factor distorsionador de las TIP ni de los juicios, sino que correlacionó positiva y significativamente con ambos y con las covariaciones reales entre las categorías. En un nivel individualizado de análisis, se encontró una variabilidad interindividual significativa en todas las posibles asociaciones de las variables estudiadas, destacando las diferencias entre los sujetos en cuanto al grado de influencia de la TIP sobre los juicios retrospectivos, y el componente semántico de los juicios.

Palabras clave: Teorías implícitas de la personalidad.

Validity of teacher's implicit personality theories

Abstract

Four analogous teacher samples helped to test the hypotheses about the validity of their implicit personality theories (IPTs). The standard of validity was based on judgments made immediately following the observation of video recorded behavior sequences. In an aggregated level of analysis, IPTs correlated with the real covariations among categories, demonstrating their structural fidelity. IPTs also contributed to the validity of social judgments, too. On the other hand, semantics did not distort either IPTs or judgments, but correlated positively and significantly with both, and with real covariations among categories. In an individualized level of analysis, there was found a significant interindividual variability in all possible associations among the variables studied, mainly the influence of IPT on retrospective judgments, and the semantic component of judgments.

Key words: Implicit personality theories.

Agradecimientos: Esta investigación ha sido financiada con una beca predoctoral del Programa Sectorial de Becas de Formación del Profesorado y Personal Investigador en España (Dirección General de Investigación Científica y Técnica, Ministerio de Educación y Ciencia).

Dirección del autor: Escuela Superior de CC. de la familia, Universidad Pontificia de Salamanca, C/ Compañía, 5, 37008 Salamanca

INTRODUCCION

El rasgo de personalidad ha demostrado ser un constructo o categoría comúnmente utilizada por los individuos para describir a las personas de su ámbito social (Hampson, 1983; Winter y col., 1985; Park, 1986). Pero este constructo no es inferido por los sujetos basándose únicamente en los datos de la «realidad» (consistencia de la conducta), sino que su uso se apoya también en el conocimiento previo. Entre este conocimiento previo se encuentran las teorías implícitas de personalidad (TIP), esquemas mediante los que los perceptores suponen correspondencias o relaciones entre las conductas y los rasgos (Cantor y Mischel, 1979a; Hampson, 1982; Buss y Craik, 1983), y entre los rasgos mismos (Bruner y Tagiuri, 1954; Rosenberg y Sediak, 1972; Schneider, 1973). El presente artículo describe una investigación en que se comprueba si esta forma de conocimiento previo se corresponde con los datos de la realidad. Dicha correspondencia la hemos sometido a prueba en el ámbito escolar, en el que las percepciones docentes sobre la personalidad infantil tienen una especial relevancia para el desarrollo del rendimiento, autoconcepto y personalidad de los niños.

EL PODER SESGANTE DE LA TIP COMO FORMA DE CONOCIMIENTO PREVIO

La TIP es entendida como una estructura esquemática multidimensional de conocimiento previo que es activada cuando el perceptor realiza, espontánea o intencionalmente, inferencias o juicios sociales que incluyen conductas y rasgos de personalidad. La investigación de enfoque estructural se ha ocupado de identificar las dimensiones de personalidad a lo largo de las cuales los sujetos clasifican a las personas-estímulo de su entorno social (Rosenberg y col., 1968; Rosenberg y Sediak, 1972a; Rosenberg y Jones, 1972; Ebbesen y Allen, 1979; Swede y Tetlock, 1986; Smith y Kihlstrom, 1987; Borkenau, 1988; Rosenberg, 1989).

Aunque el apoyo en este esquema implícito supone un notable ahorro de esfuerzo, y confiere al perceptor un sentido de control de su entorno, también puede conducir a errores perceptivos. De hecho, ya existe evidencia que demuestra que varios tipos de esquemas y formas de conocimiento previo sesgan con frecuencia los procesos perceptivos, e introducen error en los juicios sociales. Así lo ha comprobado la línea de investigación, iniciada por Snyder y Swann (1978), sobre las estrategias en la atención y selección de la información cuando el perceptor trata de comprobar una hipótesis previa. También lo confirma la línea iniciada por Hatie y Kumar (1979) que, utilizando un paradigma de formación de impresiones, demuestra los efectos del conocimiento previo sobre la recuperación de la información. Los efectos de este conocimiento también se han mostrado en áreas de investigación más concretas, como por ejemplo, en la investigación cognitiva sobre prototipos (p. ej., Cantor y Mischel, 1979b; Mayer y Bower, 1986), sobre estereotipos (p. ej., Hamilton y Rose, 1980; Kim y Baron, 1988) y sobre el yo (para revisión, cfr., Markus y Wurf, 1987; Berkowitz, 1988).

Hay que matizar que los sesgos y errores en los procesos y productos perceptivos no son sólo provocados por las TIP y demás esquemas implícitos, sino también por otros factores cognitivos y motivacionales (para revisión, cfr. Nisbett y Ross, 1980; Kruglanski y Ajzen, 1983).

LA VALIDEZ DE LA TIP COMO FIDELIDAD ESTRUCTURAL

Si bien la TIP influye en la validez de los juicios de rasgos aislados, también puede estudiarse la validez de la propia TIP, es decir, la correspondencia de las covariaciones empíricas de los rasgos, o de las conductas y los rasgos, con las covariaciones percibidas o estimadas. Este segundo tipo de validez ha sido denominado «fidelidad estructural» (Loevinger, 1957). Por tanto, hay que distinguir entre (1) la fidelidad estructural de la TIP, y (2) la validez de los juicios sociales aislados, en la que incide, entre otros factores, la TIP del perceptor.

La fidelidad estructural de las relaciones percibidas entre los rasgos, y entre las conductas y los rasgos, puede obtener la validación mediante su contraste con las relaciones reales entre los rasgos, y entre las conductas y los rasgos. Las covariaciones empíricas de rasgos (o de conductas-rasgos), que constituyen el criterio de validez, se inferirían a partir de los juicios verbales realizados inmediatamente después de la observación de la conducta o bien a partir de juicios de autoinforme. Este procedimiento empírico para dilucidar la fidelidad estructural de las TIP ha sido utilizado por Jackson y otros autores (Lay y Jackson, 1969; Lay y col., 1973; Stricker y col., 1974; Reed y Jackson, 1975; Jackson y col., 1979; Chan y Jackson, 1979, 1982). Un segundo procedimiento empírico para validar las TIP ha consistido en la comparación de la covariación estimada con la covariación conceptual (Shweder y D'Andrade, 1979, 1980; Shweder, 1975, 1982). Otra perspectiva ha abordado el estudio de la fidelidad estructural de las TIP comparando las teorías profanas con las teorías científicas (Sternberg y col., 1981; Semin y col., 1981; Furnham, 1984; Semin y Chassein, 1985; Sternberg, 1985; Semin y Krahe, 1987).

Además de los procedimientos anteriores, que pueden considerarse directos, existirían al menos otras tres líneas de investigación de cuyos resultados se desprenden interpretaciones relevantes para la cuestión de la validez de las TIP: (1) los estudios sobre la capacidad humana para detectar covariación entre eventos sociales (para revisión, cfr. Crocker, 1981), (2) las investigaciones sobre el origen de las TIP (Lewicki, 1982; Mayer y Bower, 1986), y (3) los estudios que han comprobado el nivel de consenso entre los perceptores, bien sea comparando heteroinformes con autoinformes en la evaluación de personas extrañas (Passini y Norman, 1966; Albright y col., 1988; Watson, 1989), o bien comparando los juicios de varios perceptores sobre una persona-estímulo con los juicios de un único perceptor sobre varias personas-estímulo (Dornbusch y col., 1965; Park, 1986).

La hipótesis de la solapación sistemática

A pesar del sesgo sistemático contenido en las TIP, que ha sido demostrado por el segundo tipo de procedimientos directos, y a pesar de otros sesgos presentes en la formación y desarrollo de conceptos correlacionales (Nisbett y Ross, 1980; Crocker, 1981), se ha encontrado una moderada validez correlacional (fidelidad estructural) en las TIP, tal como lo indican los estudios que han seguido los procedimientos directos 1.º y 3.º, y los procedimientos indirectos 2.º y 3.º.

Las investigaciones de Borkenau y Ostendorf (1987a, 1987b; Borkenau, 1986a, 1986b) y su *hipótesis de la solapación sistemática* constituyen un puente entre la hipótesis de la distorsión sistemática de Shweder y D'Andrade (1979, 1980; Shweder, 1975, 1982), que negaba la validez de las TIP, y la hipótesis

de la validez de las TIP, defendida por Jackson y otros autores (Lay y Jackson, 1969; Lay y col., 1973; Stricker y col., 1974; Reed y Jackson, 1975; Jackson y col., 1979; Chan y Jackson, 1979, 1982). Borkenau y Ostendorf (1987b) midieron las TIP pidiendo a los sujetos que realizaran estimaciones del tipo «Si un sujeto exhibe la conducta A frecuentemente, ¿qué probabilidad hay de que exhiba la conducta B frecuentemente?». Se encontró que los sujetos efectuaron estimaciones completamente simétricas de las probabilidades condicionales, aun cuando conocían las frecuencias-base de las conductas. Por ejemplo, aunque sabían que la conducta A era más frecuente que la B, no estimaron que la probabilidad de A, dada B, fuera mayor que la probabilidad de B, dada A, como hubiese sido lógico, sino que los juicios sobre ambas probabilidades fueron parecidos. La lógica indica que la probabilidad condicional de la actividad más frecuente debe ser más alta. Las TIP fueron comparadas no con datos de autoinforme como en los estudios anteriores, sino con las intercorrelaciones de frecuencias de conductas estimadas retrospectivamente, así como con las intercorrelaciones de frecuencias de conducta estimadas inmediatamente después de la observación de la conducta real. También se averiguó su correlación con las relaciones semánticas entre las conductas. De esta forma, se establecen las tres matrices correlacionales en las que se basó la hipótesis de la distorsión sistemática, además de las estimaciones de probabilidad condicional.

Todos los coeficientes de correlación calculados por Borkenau y Ostendorf tuvieron un signo positivo, pero el mayor fue el encontrado entre las TIP y las relaciones semánticas. Podría concluirse que los jueces estimaron las probabilidades condicionales principalmente sobre una base semántica, sin ser conscientes de que la información-base sobre la frecuencia de ocurrencia de las conductas es un requisito para una estimación precisa de las probabilidades condicionales, violando así las reglas del razonamiento formal y matemático (Kahneman y col., 1982). Sin embargo, los autores también encontraron correlaciones altas entre las estimaciones de probabilidad condicional y las de frecuencia de conductas, tanto inmediatas como retrospectivas. Esto sugeriría una cierta precisión o fidelidad estructural de las TIP en un sentido correlacional, por lo que se confirman los hallazgos optimistas de Jackson y otros autores acerca de la naturaleza realista de las teorías implícitas. Además de esta confirmación, Borkenau y Ostendorf demuestran su hipótesis de la solapación sistemática, según la cual las correlaciones entre evaluaciones retrospectivas de frecuencias de actos son explicadas, en gran parte, por la solapación semántica de las categorías usadas. Cuanto mayor es la semejanza semántica entre dos términos descriptivos de conductas, mayor es su pertenencia a características de personalidad que se solapan y, por tanto, a universos solapados de actos.

LA VALIDEZ DE LA TIP DE LOS PROFESORES

La investigación que se va a describir en este informe replica, amplía y contextualizada la comprobación de la hipótesis de la solapación sistemática de Borkenau y Ostendorf (1987a, 1987b; Borkenau, 1986a, 1986b), permitiendo demostrar la validez de las TIP. Nuestra investigación persigue objetivos semejantes a los de Borkenau y Ostendorf (1987b), aunque algunas condiciones metodológicas han sido cambiadas; se amplía el análisis de los datos al añadir un nivel individualizado al análisis que realizaron Borkenau y Ostendorf; y por úl-

timo, se contextualiza la comprobación de validez de las TIP al situarnos en el entorno escolar y estudiar la percepción de sujetos con un rol concreto en este entorno: los profesores.

En el contexto escolar, la expectativas y juicios del profesor sobre la conducta del alumno tienen su origen, en buena parte, en las TIP del profesor. La mayor parte de los estudios estructurales sobre TIP de los profesores han arrojado entre tres y cinco dimensiones a lo largo de las cuales los profesores evalúan la personalidad de sus alumnos (p. ej., Huber y Mandl, 1979; Tsoi y Nicholson, 1982; Digman —en Digman y Takemoto— Chok, 1981; Digman e Inouye, 1986). La mayoría de las dimensiones son comunes a la mayoría de los estudios. No obstante, la metodología predominante no ha permitido hallar diferencias individuales entre la TIP docentes, aunque estas diferencias probablemente existan.

La importancia del estudio de la validez de las TIP de los profesores viene dada por el hecho de que la percepción social del profesor constituye una situación de conocimiento interpersonal que tiene una especial repercusión en la conducta de la persona-estímulo. Así lo ha demostrado la línea de investigación sobre los efectos de las expectativas docentes (Rosenthal y Jacobson, 1968. Para revisión, cfr. Brophy, 1983; Miller y Turnbull, 1986). Si seguimos interpretando la evidencia hallada, podríamos afirmar que una percepción sagaz, basada en una TIP con una notable fidelidad estructural, es más beneficiosa para los aprendizajes, el autoconcepto y, en general, la conducta de los niños, que una percepción azarosa o sesgada por una TIP con una fidelidad estructural pobre. Esta conclusión se desprende de los estudios que, por ejemplo, han encontrado una potente covariación percibida entre atractivo físico y sociabilidad (Kenealy y col., 1988), entre sexo y actitudes hacia la escuela (Darom y Rich, 1988), o entre origen étnico y rendimiento (Guttman y Bar-Tal, 1982). No resulta difícil que si los profesores mantienen con perseverancia creencias covariacionales de este estilo, éstas se autocumplan y, por ejemplo, los niños menos favorecidos físicamente lleguen a ser los menos sociables, las actitudes de los niños hacia la escuela sean más negativas que las de las niñas, y los niños pertenecientes a grupos étnicos minoritarios lleguen a alcanzar un rendimiento inferior al de los niños del grupo social mayoritario. Lo que podría ser una creencia covariacional implícita y funcional del profesor, se habría convertido en una conducta disfuncional del alumno. Aun en el caso de que la creencia tuviera una base real, el perceptor exagera a menudo la covariación, provocando su acentuación en la relación real de los atributos discentes.

NUESTRO ESTUDIO

Todas las consideraciones anteriores justifican la necesidad de estudiar la fidelidad estructural de las TIP de los profesores, así como la medida en que las TIP contribuyen a la realización de juicios sagaces. En definitiva, el objetivo general que perseguimos en la presente investigación fue obtener evidencia empírica sobre la validez de las TIP de los profesores de EGB. Mientras que en la literatura hallamos estudios sobre la validez de los juicios de los profesores (p. ej., Jackson, 1968; Connolly y Doyle, 1981; Roberts y Jenkins, 1982; French y Waas, 1985; Itskowitz y col., 1988; Granleese y col., 1989), investigaciones estructurales sobre las TIP docentes (p. ej., Huber y Mandl, 1979; Tsoi y Nicholson, 1982; Digman —en Digman y Takemoto— Chok, 1981; Digman e Inou-

ye, 1986), e incluso estudios sobre la validez de las TIP de los alumnos (Whitely y Doyle, 1976; Larson, 1979; Cadwell y Jenkins, 1985) no encontramos ninguna investigación sobre la validez de las TIP de los profesores. Esta tarea fue acometida en la investigación que describimos en este artículo.

Nuestro estudio comprobó la fidelidad estructural de las TIP y la contribución de éstas a la precisión de los juicios sociales. Por otra parte, se examinó la relación de la semántica con las TIP, puesto que a menudo se ha atribuido a la semántica un papel distorsionador en la percepción de covariaciones entre rasgos. Por último se analizó si existen diferencias individuales en cuanto a la sagacidad perceptiva, la fidelidad estructural de las TIP, su utilización en la emisión de juicios sociales, y la incidencia de la semántica en las TIP y en los juicios retrospectivos. El estudio se basó fundamentalmente, tanto a nivel conceptual como metodológico, en los trabajos de Borkenau y Ostendorf (1987a, 1987b; Borkenau, 1986a, 1986b), y de forma indirecta, en los de Shweder y D'Andrade (1979, 1980; Shweder, 1975, 1982), y en los de Jackson y otros autores, que defienden la teoría de la validez de las TIP (Lay y Jackson, 1969; Lay y col., 1973; Stricker y col., 1974; Reed y Jackson, 1975; Jackson y col., 1979; Chan y Jackson, 1979, 1982). La originalidad reside, en primer lugar, en la población que hemos tomado como objeto de estudio. Los resultados obtenidos fueron interpretados dentro del marco de la investigación existente sobre la validez de las percepciones docentes. Asimismo, existe un aspecto metodológico de carácter innovador; el análisis de la validez de las TIP. Este se efectúa a dos niveles: (1) agregando los datos a través de los sujetos y (2), tomando los datos individualizadamente. Este último nivel permite conocer la variabilidad interindividual de la validez de las TIP, y no fue considerado por las tres líneas de investigación en las que, directa o indirectamente, se basa nuestro estudio.

La hipótesis general que planteamos es doble: (1) la percepción de covariación de categorías por parte de los profesores de EGB para describir la personalidad de sus alumnos, corresponde con la covariación real de las conductas, y además (2) la covariación percibida de tipo implícito contribuye a la validez de los juicios interpersonales en el aula. Esta validez fundamental de las TIP de los profesores vendrá demostrada por el hallazgo de coeficientes de correlación positivos y distintos de cero, es decir, significativos ($\alpha = 0.01$). Esta hipótesis doble es susceptible de desglosamiento en otras más particulares:

1. La estructura de la ocurrencia real de las categorías estará asociada con la TIP. Esto significa que la TIP posee una considerable validez en un sentido correlacional.
2. La estructura correlacional de los juicios retrospectivos reflejará la estructura correlacional que constituye la TIP, lo que significa que la TIP está relacionada con la emisión de juicios sociales.
3. La correlación entre la ocurrencia real de las categorías y su ocurrencia estimada indicará un notable grado de precisión estereotípica de los juicios retrospectivos (Cronbach, 1955). Si se cumple esta predicción, junto con la del apartado anterior, se podría inferir en un sentido débil que la TIP contribuye en parte a la validez de los juicios de rasgos aislados.
4. La estructura correlacional de los juicios retrospectivos reflejará también la estructura correlacional presente en las conductas reales.
5. Las relaciones semánticas entre las categorías estarán asociadas con la coocurrencia percibida de éstas (TIP).

6. Las relaciones semánticas entre las categorías-rasgo y categorías-conducta también estarán asociadas con la estructura de la ocurrencia real de categorías.
7. Por último, las relaciones semánticas entre las categorías estarán asociadas con la estructura de los juicios retrospectivos de frecuencias de las categorías.

La confirmación de las tres últimas hipótesis demostrará que la semántica no distorsiona las relaciones entre las categorías presentes en las TIP, ni la influencia de éstas en la emisión de juicios sociales precisos. Las siete hipótesis están avaladas por las investigaciones de Borkenau y Ostendorf (1987a, 1987b; Borkenau, 1986a, 1986b). Ahora bien, todas estas hipótesis se deben comprobar con un análisis agregado de los datos, mientras que, como ya se ha mencionado, también nos interesa un tipo individualizado de análisis. De acuerdo con éste se comprobarán cinco hipótesis adicionales:

8. Se encontrarán diferencias interindividuales en la sagacidad de los juicios retrospectivos (en esta hipótesis la sagacidad se entiende como la capacidad para detectar intercorrelaciones entre las categorías).
9. Se encontrarán diferencias interindividuales en el grado de asociación entre la semejanza semántica de las categorías y las TIP individuales.
10. Se encontrarán diferencias interindividuales en el grado de relación entre la semejanza semántica de las categorías y las estructuras individuales de los juicios retrospectivos de frecuencias de las categorías. Esta hipótesis está avalada por el estudio de Lamiell y col. (1980).
11. Existirán algunas diferencias entre los sujetos en cuanto a la fidelidad estructural de las TIP.
12. Existirán algunas diferencias entre los sujetos en cuanto a la utilización de las TIP en la emisión de juicios sociales.

En las siete hipótesis primeras se predice el hallazgo de coeficientes de correlación positivos y distintos de cero, es decir, significativos ($\alpha = 0.01$). En las cinco hipótesis últimas se predice el hallazgo de una desviación típica significativa en la distribución de correlaciones de cada hipótesis ($\alpha = 0.01$).

METODO

Muestra

Ante la dificultad de reclutar muestras naturales de profesores para este estudio, utilizamos muestras análogas, que estuvieron formadas por estudiantes universitarios de Pedagogía que ya poseían el título de Profesor de EGB. Al menos, estos sujetos ya poseían una experiencia docente mínima, adquirida durante su período de formación. Por otra parte, era probable que estos sujetos se dedicaran en el futuro a la enseñanza, lo que también justificó su inclusión en las muestras.

El procedimiento general de muestreo no fue probabilístico. Se invitó a participar en este estudio a los sujetos de la Universidad Pontificia de Salamanca que poseían las características antes mencionadas. Aceptaron participar 67 sujetos, aunque por dificultades prácticas no pudieron ser asignados al azar a las cuatro muestras que se necesitaban. La distribución quedó de la siguiente manera:

1. ^a muestra (Preestudio)	16 sujetos
2. ^a muestra (Valoraciones de prototipicismo)	18 sujetos
3. ^a muestra (Juicios retrospectivos de frecuencia y TIP)	17 sujetos
4. ^a muestra (Juicios de semejanza semántica)	17 sujetos

Medida y procedimiento

Preestudio: Descripción de conductas observadas. El objetivo de este preestudio era descubrir un conjunto de categorías representativas del lenguaje natural que los profesores de EGB utilizan para describir a los niños. Estas categorías servirían para elaborar posteriormente los cuatro instrumentos de medida que se emplearon en el resto del estudio. Los 16 sujetos de una primera muestra observaron, en una grabación de vídeo de veinte minutos, las actividades desplegadas por cinco niños (personas-estímulo) en un contexto extraescolar artificial (taller de dramatización creado en el laboratorio). Aunque el contexto era artificial, este tipo de material estimular es más realista y natural que los materiales utilizados en la mayor parte de los estudios sobre los TIP, en que se suelen emplear estímulos verbales escritos. En el presente preestudio se observaron directamente conductas en personas-estímulo reales.

En la grabación se intentó que la actividad desarrollada por los niños permitiese la interacción frecuente, ya que un contexto interactivo favorece mucho más la atribución de rasgos. Para ello, una colaboradora del experimentador, con titulación de Profesora de EGB, organizó y dirigió una actividad de dramatización, y estimuló a los niños para que cooperasen entre sí.

La unidad de observación no estuvo constituida por una conducta concreta, ni tampoco interesó la secuenciación temporal, sino que lo importante fue el registro narrativo del todo (Anguera, 1983). Se trataba de obtener un registro no sistematizado que permitiese elaborar un código molecular de categorías que fuera útil para las observaciones sistematizadas posteriores. Cada dos minutos se detuvo el vídeo y se pidió a cada uno de los sujetos que describiese durante otros dos minutos las conductas de una de las personas-estímulo, aleatoriamente asignada. Este muestreo focal para cada perceptor se realizó para eliminar su dispersión perceptiva y favorecer su atención selectiva. El investigador estimuló a los sujetos para que no se limitaran a realizar descripciones superficiales. Esta estimulación es necesaria para conseguir juicios naturales (Rosenberg, 1989). El plan de la sesión fue el siguiente;

- 1.º Información a los sujetos sobre el objetivo general de la sesión, y sobre la edad aproximada (10 años) de los niños que van a describir.
- 2.º Asignación aleatoria de las personas-estímulo, una por cada sujeto.
- 3.º Explicación del procedimiento (intervalos de dos minutos de observación seguidos de intervalos de dos minutos de descripción) y de la necesidad de realizar descripciones exhaustivas.
- 4.º Ejecución del procedimiento.

Los términos-conducta empleados en las descripciones se tradujeron en un sistema molecular de categorías, utilizando un conjunto de reglas codificadoras. Fueron extraídas 307 categorías de las descripciones de los 16 sujetos. De acuerdo con el objetivo del preestudio, se seleccionó un conjunto representativo de categorías según la frecuencia de aparición en las descripciones, y en un número no demasiado grande para no incrementar excesivamente el número de ítems

de los cuestionarios de posteriores medidas. Según estos criterios, se eligieron todas las categorías empleadas por cuatro o más sujetos ($n = 18$).

Valoraciones de prototipicismo. El objetivo de esta medida era descubrir las relaciones reales entre las categorías, para lo cual se realizaron juicios en una situación de observación sistemática (inmediatamente después de la observación de cada secuencia de conductas). Se entregó a cada uno de los 18 sujetos de la segunda muestra un cuestionario en el que en primer lugar se explicaba la tarea, y posteriormente aparecía una única hoja de respuesta encabezada por el nombre de una de las cinco personas-estímulo, previamente asignada de forma aleatoria a cada juez. Por tanto, tres personas-estímulo fueron juzgadas por cuatro sujetos cada una, y las dos restantes, por tres sujetos. Debajo del nombre de la persona-estímulo se encontraba una tabla de doble entrada con las categorías seleccionadas en el preestudio exploratorio situadas en las filas, y los intervalos de dos minutos cada uno, situados en las columnas. El orden de las categorías fue aleatorizado para cada sujeto que juzgaba a la misma persona-estímulo. El fraccionamiento de las actividades exhibidas en la grabación por las personas-estímulo fue el mismo que el realizado en el preestudio.

Se mostró a los sujetos cada secuencia de dos minutos, y a continuación se les pidió que evaluaran en una escala de 7 puntos (+3 = ejemplo extremadamente bueno de la categoría; -3 = contra-ejemplo extremo de la categoría; 0 = conducta irrelevante para la categoría) el prototipicismo de las actividades observadas respecto de las categorías en la hoja de la persona-estímulo, siempre que esta persona hubiera intervenido en la secuencia correspondiente. La tarea fue precedida de un ejemplo mostrado por el experimentador.

Juicios retrospectivos de frecuencia. El objetivo de esta medida era obtener juicios después de haber transcurrido un lapso de tiempo desde la observación de las conductas. Los 17 sujetos de la tercera muestra observaron las actividades de las cinco personas-estímulo, después de haberles explicado la naturaleza de la tarea evaluadora que seguiría a la observación del vídeo. Para esta tarea se empleó un cuestionario que los sujetos recibieron después de la observación del vídeo. El cuestionario constaba de una hoja por cada categoría-conducta seleccionada en el preestudio, que contenía el nombre de esa categoría y una lista de las cinco personas-estímulo (el orden de las personas-estímulo fue aleatorizado para cada categoría). Los sujetos fueron instruidos para indicar con qué frecuencia cada persona-estímulo había actuado en la forma descrita por cada categoría. La valoración de la frecuencia se efectuó en una escala de siete puntos (0 = nunca; 6 = continuamente). La tarea fue precedida de un ejemplo mostrado por el experimentador.

Medición de las teorías implícitas de personalidad. Las categorías extraídas en el preestudio fueron insertadas en los ítems para medir la TIP. Los ítems tuvieron el siguiente formato en el caso de que las categorías fueran rasgos de personalidad: «Si un alumno es (rasgo X), ¿qué probabilidad hay de que también sea (rasgo Y)?» Cuando se trataba de categorías de conducta, los ítems fueron del tipo «Si un alumno (conducta X), ¿qué probabilidad hay de que también (conducta Y)?» (las categorías-rasgo están descritas con un adjetivo, y las categorías-conducta están descritas con un verbo). Un tercer tipo de ítems contenían en una proposición una categoría-rasgo, y en otra, una categoría-conducta.

La primera hoja del cuestionario explicaba a los sujetos la naturaleza de la tarea. Las estimaciones de probabilidad condicional se realizaron en una escala

de siete puntos, siendo 1 «muy improbable» y 7, «muy probable». El número de ítems fue de 306. El orden de los ítems fue aleatorizado en una primera versión del cuestionario. Como no se disponía de un programa informático para aleatorizar el orden de los ítems para cada sujeto, la primera versión aleatorizada del cuestionario se dividió en cuatro partes. La combinación de estas cuatro partes dio lugar a cuatro cuestionarios con una ordenación diferente de los ítems. En el análisis se incluyó la comprobación de posibles efectos de orden de los ítems.

Aunque no es un objetivo de este estudio averiguar el grado de simetría en las estimaciones de probabilidad condicional, interesando sólo conocer la precisión de la TIP en un sentido correlacional, se aprovechó el formato de los ítems para dar un índice de esta simetría.

La muestra que se utilizó aquí fue la misma que la empleada para los juicios retrospectivos de frecuencias ($N = 17$), aunque la medición de la TIP se realizó una semana después de la de los juicios retrospectivos.

Juicios de semejanza semántica. Una cuarta muestra de 17 sujetos evaluó las relaciones de semejanza semántica entre las categorías. El n.º de ítems fue de 153. La primera hoja del cuestionario explicaba a los sujetos la naturaleza de la tarea. Las valoraciones se efectuaron en una escala de 7 puntos (+ 3 = sinónimos; -3 = antónimos; 0 = ausencia de relación semántica). El orden de presentación de los ítems fue aleatorizado en una primera versión del cuestionario. Posteriormente, se procedió de la misma forma que con el cuestionario anterior (Medición de las TIP), para poder descubrir posibles efectos de orden de los ítems.

RESULTADOS

Se recogieron 12.420 juicios entre las cuatro medidas. Reseñamos en primer lugar los resultados de los análisis previos para cada medida, y posteriormente, describiremos los resultados relevantes para la comprobación de las hipótesis.

Juicios retrospectivos de frecuencia. En primer lugar se calculó la fiabilidad de los juicios. Se utilizaron correlaciones intraclase, según la taxonomía que propusieron Shroul y Fleiss (1979) para las diferentes situaciones de juicio. Para este cálculo se realizó un análisis de varianza (ANOVA) bidireccional para cada una de las 18 categorías, tanto para los jueces individuales (fiabilidad teórica de las puntuaciones de cada juez) como para las estimaciones medias de los 17 jueces. Por tanto, se calcularon 36 coeficientes de fiabilidad. La fiabilidad de las estimaciones medias es relevante para el nivel de análisis agregado de los datos (hipótesis 2, 3, 4 y 7), mientras que la fiabilidad para los jueces individuales es relevante para el análisis individualizado (hipótesis 8, 10 y 12). También se halló la significación de los coeficientes de correlación intraclase.

En la tabla 1 presentamos el análisis de fiabilidad, junto con las medias y desviaciones típicas de los juicios retrospectivos.

La media de la fiabilidad de las valoraciones medias, calculada a partir de la transformación Z de Fisher, es de 0.728, y la de la fiabilidad de las valoraciones individuales es de 0.154. Tanto la fiabilidad de las valoraciones individuales para cada categoría, que oscila entre 0.010 y 0.317, como la fiabilidad media (0.154), son bajas, si bien hemos visto que la mayor parte son significativas. En

general, los coeficientes de fiabilidad se puedan considerar suficientemente satisfactorios.

En segundo lugar, los juicios fueron promediados a través de jueces y personas-estímulo para cada categoría, tal como lo requería la hipótesis 3. Posteriormente se hallaron las intercorrelaciones entre los juicios retrospectivos para las categorías, emitidos por cada juez. Por tanto, obtuvimos 17 matrices correlacionales de 18×18 categorías (153 correlaciones en cada matriz), que denominamos JRi (el cálculo de las 17 matrices era necesario para investigar las diferencias individuales). También se calcularon las intercorrelaciones entre los juicios retrospectivos para las categorías, pero teniendo en cuenta los juicios de todos los sujetos sobre todas las personas-estímulo. La matriz resultante, de 18×18 categorías, constó de 153 coeficientes de correlación, y revelaba la estructura de los juicios sobre la frecuencia estimada de las categorías. Fue denominada JRt (el cálculo de esta matriz era necesario para el subsiguiente análisis agregado de los datos). Las correlaciones de las 17 matrices JRi y de la matriz JRt fueron transformadas en Z de Fisher. Puesto que la comprobación de las hipótesis exigía correlacionar correlaciones, la transformación de las correlaciones en puntuaciones medidas en una escala de intervalo resultaba más apropiada para los supuestos de la correlación de Pearson.

TABLA 1

Medias, desviaciones típicas, fiabilidad de las valoraciones medias y fiabilidad de las valoraciones individuales de los juicios retrospectivos de frecuencias para cada categoría

Categoría	Valoración media	Desviación típica	Fiabilidad de las valoraciones medias	Fiabilidad de las valoraciones individuales
Sociable	3.329	1.491	0.719°	0.131°
Activo/a	4.129	1.526	0.677°	0.110°
Impositivo/a	2.647	1.771	0.144	0.010
Ayuda a los demás	3.518	1.797	0.887*	0.317*
Habla con los demás	3.835	1.731	0.694°	0.118°
Se divierte	4.341	1.484	0.516	0.060
Disfruta con lo que hace	4.153	1.323	0.820*	0.211*
Introverso/a	2.765	1.925	0.550	0.067
Centrado/a en su trabajo	4.059	1.491	0.736*	0.141*
Decidido/a	3.718	1.716	0.740°	0.144°
Tranquilo/a	3.165	1.975	0.741*	0.144*
Creativo/a	3.671	1.592	0.551	0.067
Acepta las ideas que le dan	3.447	1.547	0.704°	0.123°
Acepta la ayuda de otros	3.882	1.546	0.881*	0.303*
Dependiente	3.224	1.762	0.849*	0.245*
Seguro/a de sí mismo	3.718	1.790	0.814*	0.204*
Obs. lo que hacen sus comp.	2.965	1.614	0.642°	0.095°
Pide ayuda	3.318	1.545	0.852*	0.253*

* Coeficiente significativo al nivel de confianza del 1 por ciento.

° Coeficiente significativo al nivel de confianza del 5 por ciento.

Valoraciones de prototipicismo. El primer caso de la taxonomía de Shrout y Fleiss (1979) es el que tomamos para hallar la fiabilidad de las valoraciones de prototipicismo. No nos interesaba hallar la fiabilidad de las valoraciones de los jueces individuales, ya que sólo vamos a utilizar en análisis posteriores los datos agregados a través de personas-estímulo y jueces (hipótesis 1, 3, 4, 6, 8 y 11). El modelo estadístico apropiado es el ANOVA unidireccional.

Antes de efectuar los 18 ANOVA (uno por cada categoría) y puesto que los n juicios de los grupos era desigual, se decidió verificar la homogeneidad de las varianzas de los cinco grupos, y se aplicó la prueba de homoscedasticidad de Bartlett. Se demostró que la varianzas de los grupos incluidos en cada ANOVA eran homogéneas, por lo que se procedió a la realización de los ANOVA y al cálculo de los coeficientes de correlación intraclase y su significación. En la tabla 2 se presenta la fiabilidad de las valoraciones medias, junto con las medias y desviaciones típicas de las valoraciones de prototipicismo.

Los coeficientes de fiabilidad varían mucho de una categoría a otra, oscilando entre -0.885 y 0.913 . El coeficiente de fiabilidad medio, calculado a partir de la transformación Z de Fisher, es de 0.130 . Este se puede considerar bastante bajo. Es decir, sorprendentemente se constata un acuerdo bajo entre los jueces, resultando todavía más desconcertante la presencia de 8 coeficientes con signo negativo (la media cuadrática entre los grupos fue menor que la de dentro de los grupos, en los ANOVA correspondientes).

TABLA 2

Medias, desviaciones típicas, y fiabilidad de los juicios medios, a través de los jueces, de las valoraciones de prototipicismo para cada categoría

Categoría	Valoración media	Desviación típica	Fiabilidad de las valoraciones medias
Sociable	9.944	8.861	-0.672
Activo/a	14.722	7.637	-0.329
Impositivo/a	1.833	8.291	-0.331
Ayuda a los demás	4.444	10.285	0.913*
Habla con los demás	11.944	8.025	0.464
Se divierte	11.000	9.418	-0.338
Disfruta con lo que hace	11.056	9.149	0.533
Introvertido/a	0.556	12.006	0.392
Centrado/a en su trabajo	13.000	7.276	0.340
Decidido/a	9.833	8.452	0.273
Tranquilo/a	11.111	8.043	-0.410
Creativo/a	9.444	8.234	-0.124
Acepta las ideas que le dan	4.667	7.113	0.867*
Acepta la ayuda de otros	6.833	7.838	0.895*
Dependiente	-2.556	11.718	0.885
Seguro/a de sí mismo	11.500	8.205	-0.448
Obs. lo que hacen sus comp.	1.833	10.019	0.078
Pide ayuda	0.611	7.709	0.100

* Coeficiente significativo al nivel de confianza del 1 por ciento.

En segundo lugar, las valoraciones de prototipismo de las actividades correspondientes a cada categoría de cada persona-estímulo fueron sumadas. Estas sumas se promediaron a través de jueces y personas-estímulo para cada categoría. Por último, las medias fueron correlacionadas con los juicios retrospectivos medios para comprobar la hipótesis 3, o lo que es lo mismo, la precisión estereotípica (Cronbach, 1955).

También se averiguó la estructura real de la ocurrencia de categorías. Para llevar a cabo esto, se sumaron las puntuaciones correspondientes a cada categoría de cada persona-estímulo, y a continuación se interrelacionaron las categorías. Así, se obtuvo una matriz de 153 correlaciones que indicaba el grado de covariación real de las categorías. Se denominó VP para posteriores análisis.

Teorías implícitas de personalidad. La fiabilidad de los juicios de probabilidad condicional (medida de TIP) se estimó utilizando dos coeficientes intraclass (ICC [2,1]; ICC [2,17]), correspondientes al segundo caso de la taxonomía de Shrout y Fleiss (1979). Un coeficiente de fiabilidad correspondía a los juicios promediados y otro, a los jueces individuales. Para el cálculo de estos dos coeficientes se realizó un ANOVA bidireccional. La fiabilidad de los juicios medios fue de 0.926 (ICC [2,17]), y la de los juicios individuales, 0.422 (ICC [2,1]). Ambos coeficientes son altamente significativos, con una probabilidad α muy por debajo de 0.01.

Además de la fiabilidad, se llevaron a cabo dos comprobaciones previas que no estaban previstas en las hipótesis: los índices de simetría y los efectos de orden de los ítems. En cuanto a los índices de simetría, se trataba de saber si los jueces realizaban implícitamente estimaciones simétricas o asimétricas de probabilidad condicional, es decir, si estiman que la probabilidad de que se manifieste la categoría A dada B es la misma que la probabilidad de B dada A. Para comprobar esto, se halló una correlación, por cada sujeto, entre las puntuaciones dadas a los ítems con las categorías situadas en un cierto orden con las puntuaciones de los ítems con el mismo par de categorías, pero situadas en orden inverso. La tabla 3 muestra el grado en que los 17 sujetos percibieron como simétricas las estimaciones de probabilidad condicional. Todos los coeficientes son altamente significativos. El promedio de estos 17 coeficientes, efectuado a través de la transformación Z de Fisher, es de 0.590. Tanto el coeficiente de cada sujeto como la correlación media muestran que, en buena parte, los sujetos perciben como simétrica la probabilidad de coocurrencia de pares de categorías. Este dato corrobora la evidencia aportada anteriormente por Mirels (1976) y por Borkenau y Ostendorf (1987b).

No obstante, hasta aquí no hemos demostrado que las categorías posean diferentes frecuencias de ocurrencia, por lo que no podemos saber si la simetría percibida es sagaz o no. Podemos comenzar por observar la tabla 1: las medias de los juicios para cada categoría no son iguales, al menos aparentemente. La categoría «impositivo/a» es la que posee una frecuencia juzgada de ocurrencia más baja ($M_1 = 2.647$), mientras que la categoría «Se divierte» es juzgada como la más frecuente ($M_2 = 4.341$). Si comprobamos la hipótesis $m_1 = m_2$, siendo $H_0: m_1 = m_2$, y utilizamos el estadístico de contraste t para muestras dependientes (Glass y Stanley, 1986), encontramos que $t = 10.47 > 2.639$ ($t_{0.995,80}$) (hemos tomado 80 grados de libertad en lugar de 84 ($n-1 = 84$) por no contener la tabla utilizada el valor crítico correspondiente a 84 g. l.). Es decir, las dos medias difieren estadísticamente al nivel de confianza del 1 por

ciento. Este dato indica que los jueces han percibido distintas frecuencias de ocurrencia de las categorías «Impositivo/a» y «Se divierte». Lo mismo se podría demostrar respecto de otros pares de categorías. De hecho, en cuanto se constate $M_m - M_n < 0.427$, se confirmará H_1 y se rechazará H_0 (n. c. 1 por ciento). Si se echa un vistazo a la tabla 1 y se efectúan varias restas entre categorías, rápidamente se concluye que un grupo muy numeroso de ellas sobrepasan el resultado 0.427. Por tanto, si los sujetos juzgan que las frecuencias con las que se han manifestado las categorías son diferentes, y además son sagaces en la percepción de las diferencias de frecuencia de unas categorías respecto de otras (como más adelante se demostrará —hipótesis 3—), parecería lógico que las TIP de estos sujetos introdujeran la asimetría en sus estimaciones de probabilidad condicional. No obstante, los coeficientes de correlación que reflejan el grado de esta simetría son, como ya se ha visto, notables. Por consiguiente, las TIP podrán ser válidas en un sentido correlacional, pero no lo son si también se introduce en el criterio de validez el grado de simetría en las estimaciones de probabilidad condicional. Los jueces ignoran en alguna medida que, dadas dos categorías con diferentes frecuencias de ocurrencia ($f_A > f_B$), la probabilidad condicional de la categoría más frecuente debería ser más alta ($p [A-B], > p [B-A]$).

El último análisis previo realizado se refiere a la comprobación de la existencia de posibles efectos de orden de los ítems. Para ello, en primer lugar se promediaron las puntuaciones en cada ítem de los sujetos que habían contesta-

TABLA 3

Coefficientes de correlación indicadores del grado de simetría percibida en las estimaciones de probabilidad condicional

Sujetos	Simetría
JMA	0.561*
RMF	0.603*
EP	0.421*
PVL	0.649*
MEA	0.577*
MCG	0.573*
JMC	0.453*
LR	0.560*
JSO	0.833*
CAS	0.589*
MCA	0.598*
TPM	0.538*
MFH	0.452*
CSS	0.591*
IMB	0.729*
RMM	0.443*
MEB	0.497*

* Coeficiente significativo al nivel de confianza del 1 por ciento.

do a un mismo tipo de cuestionario. Como hay cuatro tipos de cuestionarios según el orden de presentación de los ítems, habrá cuatro grupos de puntuaciones promediadas. A continuación se correlacionaron entre sí las puntuaciones medias de los ítems de cada tipo de cuestionario. Este análisis arrojó seis coeficientes elevados y semejantes (los seis oscilaron entre 0.722 y 0.788). Es decir, el acuerdo entre los sujetos fue alto, y aparentemente no dependió de la ordenación de los ítems dentro del cuestionario. Para confirmar esto con más rotundidad, realizamos una prueba de significación de la diferencia entre la correlación más alta ($r = 0.788$) y la más baja ($r = 0.722$), con muestras dependientes. El valor de z en esta prueba fue: $z = 2.56 < 2.58 (z_{0.995})$, por lo que se confirma que no existe diferencia entre los dos coeficientes de correlación al nivel de confianza del 1 por ciento, ni tampoco existe entre ninguno de los 14 pares restantes posibles de las seis correlaciones de la tabla. En definitiva, concluimos que la ordenación de los ítems dentro de los cuestionarios no ha ejercido ningún efecto sobre las respuestas.

Posteriormente, se promediaron las respuestas a los ítems con el mismo par de categorías, por lo que obtuvimos 17 matrices de 153 puntuaciones de covariación, una por cada juez (el cálculo de las 17 matrices es necesario para investigar las diferencias individuales). Estas matrices se denominaron Tli. A continuación, las puntuaciones se promediaron de nuevo a través de los jueces. Es decir, se obtuvo una única matriz con 153 puntuaciones medias de covariación percibida entre las categorías. Esta matriz se denominó Tlt.

Semejanza semántica. Del análisis de los 17 cuestionarios de semejanza semántica se eliminaron los datos de un sujeto, debido al patrón sistemático de respuestas encontrado (140 de las 153 respuestas se situaron en el punto medio de la escala). Por tanto, se analizaron los datos de 16 cuestionarios.

La fiabilidad de los juicios medios de semejanza semántica se estimó utilizando un coeficiente intraclass (ICC [2,16]), correspondiente al segundo caso de la taxonomía de Shrout y Fleiss (1979). Para el cálculo de este coeficiente se realizó un ANOVA bidireccional. La fiabilidad de los juicios medios fue de 0.900 (ICC [2,16]), la cual es altamente significativa, con una probabilidad α muy por debajo de 0.01. No nos interesa hallar la fiabilidad de las valoraciones de los jueces individuales, ya que sólo vamos a utilizar en análisis posteriores los datos promediados a través de los jueces (hipótesis 5, 6, 7, 9 y 10).

Se comprobaron los posibles efectos de orden mediante el mismo procedimiento que en el caso de los cuestionarios para medir las TIP: verificando la significación de la diferencia entre la correlación más alta y la más baja (esta vez, con muestras independientes) entre las halladas al relacionar las puntuaciones medias de los cuatro tipos de cuestionarios. En este caso, los seis coeficientes también fueron elevados y semejantes. La prueba de significación de la diferencia entre la correlación más alta ($r = 0.758$) y la más baja ($r = 0.654$) arrojó un valor $z = 1.812 < 1.96 (z_{0.975})$, por lo que se confirma que no existe diferencia entre los dos coeficientes de correlación al nivel de confianza del 5 por ciento, ni tampoco existe entre ninguno de los 14 pares restantes posibles de las seis correlaciones de la tabla. En definitiva, concluimos que la ordenación de los ítems dentro de los cuestionarios no ejerció ningún efecto sobre las respuestas.

Por último, se averiguó la estructura real de la semejanza semántica de las categorías. Para llevar a cabo esto, se promediaron las puntuaciones de los 16

sujetos en cada uno de los 153 ítems. Así, se obtuvo una matriz de 153 puntuaciones medias de covariación semántica. Se denominó SS para posteriores análisis.

Comprobación de las hipótesis. Para poner a prueba las hipótesis, se intercorrelacionaron las matrices JRt, JRi, VP, Tlt, Tli y SS. Cuando al menos una de las dos matrices contenía correlaciones (recuérdese que las matrices Tlt, Tli y SS constan de puntuaciones de covariación que no son correlaciones) el procedimiento adoptado en el análisis agregado de los datos fue doble: (1) se correlacionaron directamente los coeficientes r de Pearson, y (2) se transformaron las correlaciones de las matrices en Z de Fisher, y se correlacionaron las matrices Z de Fisher. El segundo procedimiento se adoptó ante el temor de que no se distribuyesen normalmente las correlaciones de, al menos, una de las matrices. En el análisis individualizando de los datos sólo se siguió el segundo procedimiento, por haber comprobado en el análisis agregado que ambas estrategias arrojaban coeficientes semejantes.

La tabla 4 resume las correlaciones halladas entre todas las matrices correlacionadas.

La tabla 5 resume las correlaciones que sirven para comprobar las siete hipótesis correspondientes al análisis agregado de los datos, mientras que la tabla 6 presenta las correlaciones medias, la amplitud de estas correlaciones (en forma del coeficiente inferior y superior de la distribución), y las desviaciones típicas de las Z de Fisher en que han sido transformadas las correlaciones, todo ello referente a las cinco hipótesis correspondientes al análisis individualizado de los datos.

TABLA 4

Correlaciones entre las matrices JRt, JRi, VP, Tlt, Tli, y SS para la verificación de las hipótesis

Matrices	JRt	JRi	Tlt	Tli	VP
JRt	—				
JRi		—			
Tlt	.699		—		
Tli		-.046/.674		—	
VP	.542	-.008/.393	.551	.323/.541	—
SS	.674	.002/.581	.907	.542/.834	.542

Nota. En esta tabla sólo se presentan las correlaciones halladas mediante el procedimiento 2, es decir, transformando previamente en Z de Fisher las correlaciones de las matrices correlacionadas. Por otra parte, en el caso del análisis individualizado únicamente se presentan en la tabla la correlación más baja y la más alta de cada grupo de correlaciones.

En la tabla 5 se comprueba, en el caso de las hipótesis 1, 2, 4, 6, y 7, que los coeficientes de correlación hallados a partir de las matrices de correlaciones son casi idénticos a los calculados a partir de las matrices de Z de Fisher. En el comentario de los resultados presentados en esta tabla vamos a tomar en cuenta las correlaciones entre las matrices de Zs de Fisher.

Un primer vistazo a las tablas 5 y 6 permite ya comprobar que todas las hipótesis han sido verificadas, debido a la significación estadística de las correlaciones (hipótesis 1, 2, 3, 4, 5, 6 y 7) y de las desviaciones típicas (hipótesis 8, 9, 10, 11 y 12).

La hipótesis 3 es la única que no se basa en la correlación de matrices, sino en la correlación de las medias de dos medidas (juicios retrospectivos y valoraciones de prototipismo). El coeficiente no sólo es significativo, sino que es el segundo más alto de la tabla 5 ($r = 0.743$). Este dato permite afirmar que el juez medio percibe con precisión la frecuencia de manifestación de unas categorías respecto de otras en una persona-estímulo media. Debe notarse que la precisión estereotípica del juez medio ($r = 0.743$) no es mayor que la fidelidad

TABLA 5

Coefficiente de correlación para comprobar las siete primeras hipótesis (análisis agregado de los datos)

Hipótesis	Matrices	Coeficiente	
1 ^s	Tlt-VP	.544*	.551 ^{o*}
2	Tlt-JRt	.699*	.699 ^{o*}
4	VP-JRt	.531*	.542 ^{o*}
5	SS-Tlt	.907*	
6	SS-VP	.528*	.542 ^{o*}
7	SS-JRt	.676*	.674 ^{o*}
3 (X de jr y vp)		.743*	

^o Las correlaciones correlacionadas han sido transformadas previamente en Z de Fisher.

* Coeficiente significativo al nivel de confianza del 1 por ciento.

TABLA 6

Correlaciones medias, amplitud y desviaciones típicas para comprobar las cinco últimas hipótesis (análisis individualizado de los datos)

Hipótesis	8	9	10	11	12
Matrices	VP-JRi	SS-Tli	SS-JRi	Tli-VP	Tli-JRi
F	.274	.699	.327	.425	.296
amplitud de r	-.008/.393	.542/.834	.002/.581	.323/.541	-.046/.674
S _z	.12*	.155*	.204*	.078*	.213

* Coeficiente significativo al nivel de confianza del 1 por ciento.

estructural de la TIP ($r = 0.551$) ($z = 1.29 < 1.645$ [z 0.95]), por lo que se concluye que la percepción que tiene el juez medio sobre una persona-estímulo generalizada posee la misma sagacidad que la fidelidad con que las covariaciones percibidas entre las categorías reflejan las covariaciones reales.

El coeficiente de correlación más alto de la tabla 5 se da entre las TIP y la semejanza semántica de las categorías ($r = 0.907$), si bien no es una correlación perfecta. Se podría inferir que las TIP tienen un fortísimo componente semántico. No obstante, una correlación no nos autoriza a inferir esto con un absoluto grado de seguridad. Al contrario, también podríamos suponer que los juicios de semejanza semántica están contaminados por las TIP de los jueces. Un dato empírico que apoya esta aseveración viene dado por las bajas correlaciones entre los sujetos en sus juicios de semejanza semántica. La mayor parte de las correlaciones son del orden de 0.2..., 0.3... y 0.4..., incluso algunas de ellas no son significativas (9 de las 120 correlaciones no alcanzan significación al nivel de confianza del 1 por ciento). Si las TIP no influyeran en los juicios de semejanza semántica, el acuerdo alcanzado sería probablemente mayor (partimos de la hipótesis de que los significados lingüísticos están consensuados en una población en mayor medida de lo que lo están las covariaciones percibidas entre las categorías-rasgo y categorías-conducta). Por otra parte, cuando se analiza individualizadamente la relación de la semántica con la TIP de cada sujeto, y promediamos estas correlaciones, el coeficiente baja a 0.699. La relación de las TIP individuales con la semántica oscila entre 0.542 y 0.834, resultando significativa la variabilidad de esta relación entre los sujetos. Es decir, la semántica está más asociada a la TIP de algunos sujetos que a la de otros, aunque en general está bastante asociada a la de todos (0.542 es el coeficiente que expresa la relación más baja).

Tal como predecía la teoría previa, los juicios retrospectivos alcanzan correlaciones más altas con las TIP ($r = 0.699$) y con la semántica ($r = 0.674$), que las valoraciones de prototipicismo o los juicios inmediatos de conducta ($r = 0.551$ y $r = 0.542$, respectivamente). Dos pruebas de contraste de hipótesis para muestras dependientes revelan que $r_{TIt-JRt} > r_{TIt-VP}$, y que $r_{SS-JRt} > r_{SS-VP}$. La primera prueba arroja un estadístico $z = 2.62 > 2.33$ ($z_{0.99}$), que confirma que $r_{TIt-JRt}$ es mayor que r_{TIt-VP} . La segunda prueba arroja un estadístico $z = 2.29 > 1.645$ ($z_{0.95}$), que muestra que $r_{SS-JRt} > r_{SS-VP}$, aunque con mayor posibilidad de error α que en el caso anterior. Este patrón es lógico, puesto que los juicios retrospectivos están más influidos por las TIP y la semántica que los juicios inmediatos, al no poder aquéllos basarse tanto en las conductas reales debido al tiempo transcurrido desde la observación de la conducta, y la consiguiente pérdida y distorsión mnemónica.

A su vez, la semántica posee una relación bastante mayor con las TIP que con los juicios retrospectivos: $z = 5.78 > 2.33$ ($z_{0.99}$). Es comprensible que un juicio de covariación sobre un otro generalizado tome como referente, en buena parte, los significados abstractos de las categorías, mientras que los juicios retrospectivos conservan un referente concreto, aunque más o menos lejano, en las conductas de las personas-estímulo. Por último, como se desprende inmediatamente de los tres contrastes anteriores, la semántica está mucho más asociada a las TIP que a las valoraciones de prototipicismo: $z = 6.71 > 2.33$ ($z_{0.99}$). Por tanto, existe una jerarquía de asociación entre la semántica y las otras tres variables: (1.º) SS—TIt, (2.º) SS—JRt, y (3.º) SS—VP.

Como era lógico, los coeficientes de correlación promediados del análisis

individualizado son todos notablemente menores que los coeficientes arrojados por el análisis agregado. Las correlaciones que más coinciden en ambos niveles de análisis se refieren a la fidelidad estructural de la TIP (el porcentaje de variación común explicada es del 30.36 por ciento en el análisis agregado, y del 18.06 por ciento en el análisis individualizado). Precisamente, las hipótesis 1 y 11, junto con las 2 y 12, son las más relevantes para la hipótesis general. Pero mientras que los resultados referidos a las hipótesis 1 y 11 revelan una relativa estabilidad de la fidelidad estructural de las TIP a través de diversos tipos de análisis, los datos relativos a las hipótesis 2 y 12 no pueden ser interpretados con la misma fiabilidad. Las correlaciones correspondientes a estas hipótesis son las que menos coinciden en los dos tipos de análisis (el porcentaje de variación común explicada es del 48.86 por ciento en el análisis agregado, y del 8.76 por ciento en el análisis individualizado). Además la amplitud de las correlaciones $r_{TIt-JRt}$ de los 17 sujetos y su desviación típica son las mayores, aparentemente, entre los cinco grupos de distribuciones de correlaciones del análisis individualizado (Cfr. tabla 6), es decir, existen notables diferencias entre los sujetos en cuanto al grado de asociación entre sus TIP y la estructura de sus juicios retrospectivos.

Concretamente, si comprobamos la significación de la diferencia entre las varianzas de las distintas distribuciones de correlaciones, se observa que la varianza de la distribución $r_{TIt-JRt}$ no es mayor que la varianza de la distribución r_{SS-JRt} ($t = 0.98 > 2.602 [t_{0.99, 13}]$), pero sí lo es respecto de la varianza de la distribución r_{SS-TIt} ($t = 4.816 > 2.602 [t_{0.99, 13}]$), y por tanto, también lo es respecto de las varianzas de las distribuciones r_{VP-JRt} y r_{TIt-VP} , por ser dichas varianzas menores que la de la distribución r_{SS-TIt} . Por tanto, la variabilidad entre los sujetos en cuanto al grado de asociación entre sus TIP y la estructura de sus juicios retrospectivos no se ve superada por ninguna otra variabilidad, y sólo es comparable a la existente entre los sujetos en cuanto a la asociación de la semántica con la estructura de sus juicios retrospectivos.

En cuanto a las dos hipótesis que restan por comentar en el análisis individualizado (8 y 10), vienen demostradas por desviaciones típicas significativas, si bien la correlación media de cada distribución es notablemente más baja que las obtenidas en el análisis agregado (r_{VP-JRt} : el porcentaje de variación común explicada es del 29.38 por ciento en el análisis agregado, y del 7.51 por ciento en el análisis individualizado; r_{SS-JRt} : el porcentaje de variación común explicada es del 45.43 por ciento en el análisis agregado, y del 10.69 por ciento en el análisis individualizado). No obstante, 13 de los 17 sujetos obtienen correlaciones r_{VP-JRt} y r_{SS-JRt} significativa. 12 coeficientes r_{VP-JRt} resultaron significativos al nivel de confianza del 1 por ciento y 1, al 5 por ciento, mientras que 11 correlaciones r_{SS-JRt} fueron significativas al 1 por ciento y 2, al 5 por ciento. La variabilidad de las correlaciones r_{SS-JRt} resultó ser significativamente mayor que la de las correlaciones r_{VP-JRt} , a juzgar por una prueba de significación de diferencias entre las varianzas para muestras dependientes ($t = 2.74 > 2.602 [t_{0.99, 13}]$). Es decir, los sujetos difieren entre sí más en cuanto al componente semántico de la estructura de sus juicios que en cuanto a su capacidad para identificar retrospectivamente la estructura de la coocurrencia real de las categorías.

DISCUSION

A pesar de que la relación entre determinadas matrices es mayor que la existente entre otras, el análisis de datos agregados ha demostrado una relación sig-

nificativa, e incluso elevada, entre las valoraciones de prototipismo, las teorías implícitas de personalidad, la semántica y los juicios retrospectivos. Este hecho confirma la hipótesis general de nuestro estudio y, simultáneamente, corrobora la hipótesis de la solapación sistemática.

Los resultados del análisis agregado constituyen una réplica de los obtenidos por Borkenau y Ostendorf (1987b), al mismo tiempo que descartan la hipótesis de la distorsión sistemática de Shweder y D'Andrade (1979, 1980; Shweder, 1975, 1982), si bien afirmamos que la semántica está más relacionada con las TIP y con los juicios retrospectivos que con las relaciones reales entre las categorías. ¿Se pueden considerar nuestros resultados del análisis agregado semejantes a los de Borkenau y Ostendorf? Debido a las diferencias metodológicas de sus dos estudios y del nuestro, la comparación directa no es posible. Sin embargo, se puede realizar una comparación intuitiva; las correlaciones obtenidas por Borkenau y Ostendorf (1987b) entre las cuatro medidas semejantes a las nuestras, oscilan, en su primer estudio, entre 0.54 y 0.74 y en su segundo estudio, entre 0.43 y 0.75. En el nivel agregado de análisis de nuestro estudio, los coeficientes oscilan también entre valores semejantes a éstos (Cfr. tabla 5), excepto el que relaciona SS con TIt, que es notablemente más alto. Aun en este caso hay una coincidencia: el coeficiente más alto obtenido por Borkenau y Ostendorf en sus dos estudios es el que relaciona SS con TIt. Por tanto, aparentemente nuestros resultados son semejantes a los de Borkenau y Ostendorf (1987b).

En este estudio no se pretendía establecer el grado de influencia de las TIP sobre la semántica o de la semántica sobre las TIP. Pero sí es posible inferir que las correlaciones entre TIP por un lado, y juicios retrospectivos y valoraciones de prototipismo, por otro, pueden interpretarse como el grado de incidencia de las TIP sobre las dos últimas variables, puesto que la TIP es una estructura mnemónica previa a la realización de los juicios, inmediatos o retrospectivos.

En cuanto a la variabilidad interindividual, el análisis individualizado de los datos demuestra, en primer lugar, una dispersión significativa entre los sujetos en cuanto a (1) su capacidad para identificar retrospectivamente la estructura de la coocurrencia real de las categorías (hipótesis 8), (2) el componente semántico de sus TIP (hipótesis 9), (3) el componente semántico de la estructura de los juicios (hipótesis 10), (4) la fidelidad estructural de sus TIP (hipótesis 11), y (5) el grado de influencia de la TIP sobre los juicios retrospectivos (hipótesis 12).

Tomando conjuntamente las hipótesis 1, 11, 2 y 12, se concluirá que las TIP poseen una fidelidad estructural elevada y fiable, al mismo tiempo que poseen una influencia significativa pero moderada sobre los juicios retrospectivos, variando bastante dicha incidencia interindividualmente.

En resumen, los resultados más importantes de este estudio han sido los siguientes:

1. Las TIP, la semántica, la estructura de los juicios retrospectivos y las covariaciones reales entre las categorías se encuentran relacionadas entre sí, lo que demuestra que las TIP poseen una notable fidelidad estructural e influyen en los juicios retrospectivos, sin menoscabo de que al mismo tiempo se caracterizan por un componente semántico muy fuerte.

2. La fidelidad estructural de las TIP sólo queda demostrada en un sentido correlacional. Los elevados índices de simetría hallados en las estimaciones de probabilidad condicional de cada sujeto indican que éstos no incorporan a sus

TIP el conocimiento que tienen sobre las diferentes frecuencias de ocurrencia de las categorías.

3. Entre las variables estudiadas, la relación más fuerte se dio entre la semántica y las TIP. Por otra parte, los juicios retrospectivos demostraron tener una relación más elevada con las TIP y la semántica que las covariaciones reales entre las categorías.

4. La relación entre las TIP, la semántica, la estructura de los juicios retrospectivos y las covariaciones reales disminuye sensiblemente cuando se adopta un tipo individualizado de análisis. Los coeficientes de correlación siguen siendo significativos en su mayor parte. Sin embargo, se demuestra una amplia variabilidad entre los jueces en cuanto a la asociación de las diversas variables, llegándose a obtener incluso coeficientes de correlación no significativos.

5. La mayor variabilidad interindividual se produce en la relación entre la semántica y la estructura de los juicios retrospectivos, y en la influencia de las TIP sobre la estructura de los juicios.

6. En definitiva, las TIP de los profesores de EGB poseen una fidelidad estructural elevada y fiable, al mismo tiempo que poseen una influencia significativa pero moderada sobre los juicios retrospectivos, variando bastante dicha incidencia interindividualmente.

La fidelidad estructural de las TIP (y todas las hipótesis en cuya demostración interviene la matriz VP) se demuestra a pesar de los reducidos coeficientes de fiabilidad (la mayor parte, no significativos) que hallamos en las valoraciones de prototipicismo. La baja fiabilidad de esta medida constituye un serio inconveniente de nuestra investigación. No obstante, puesto que las correlaciones pueden corregir la atenuación debida a la ausencia de una fiabilidad mayor, la fidelidad estructural teórica de las TIP aún es mayor.

A pesar de la utilización de muestras análogas, nuestro estudio defiende la validez de una parte de la percepción docente sobre las conductas y personalidad de los alumnos. Esta defensa puede verse matizada si consideramos que el criterio con el que se confrontó la validez de las TIP docentes (valoraciones de prototipicismo efectuadas inmediatamente después de la observación de la conducta) no es perfecto, y puede estar contaminado por las propias TIP de los sujetos que efectuaron las valoraciones. La contaminación se deduciría del escaso acuerdo de los sujetos (baja fiabilidad). Este defecto del estudio puede corregirse mediante el empleo de criterios más objetivos, o mediante un procedimiento más perfecto en el uso del criterio observacional (por ejemplo, incorporar al procedimiento un cuidadoso y prolongado entrenamiento de los observadores.).

Si nos centramos en el contexto escolar, nuestra investigación es optimista en cuanto al grado de sagacidad del conocimiento docente sobre los atributos discentes. Conocida la importancia de este conocimiento, serán los aprendizajes, el autoconcepto y, en general, la conducta de los niños los que se verán favorecidos por la validez de las TIP de los profesores. No se debe olvidar, sin embargo, que las TIP no constituyen el único factor que incide en las expectativas y juicios de los profesores. Como ya se ha señalado en este artículo, existen estudios que han hallado percepciones docentes distorsionadas, y éstas pueden deberse a múltiples factores cognitivos y motivacionales.

Por otra parte, las personas-estímulo y las categorías utilizadas en los juicios son dos elementos que también influyen en los resultados. Los niños incluidos en el material-estímulo, por ejemplo, eran de raza blanca y su atractivo físico era medio. ¿Qué hubiera pasado si algunos de los niños hubiesen sido gitanos?,

¿o si hubiéramos introducido a niños obesos, hiperactivos o agresivos...? En segundo lugar, tratamos de que las categorías fueran representativas del lenguaje natural de los profesores, pero sin duda 18 categorías es un número demasiado pequeño. Futuros estudios pueden articular un diseño que incluya varias decenas de categorías, sin prolongar demasiado la longitud de los instrumentos de medida.

A pesar del reconocimiento que hemos hecho de algunos puntos del estudio que pueden limitar la validez de sus resultados (fiabilidad de una medida, criterio de validez de las TIP, personas-estímulo y categorías utilizadas), consideramos que este trabajo efectúa una aplicación válida de un área de investigación de la psicología social al estudio de la percepción docente, poniendo de relieve un objeto o esquema cognitivo (TIP) con frecuencia desestimado en la investigación educativa, tal vez por su carácter implícito.

Por último, consideramos que la investigación presente no se puede quedar en la mera formulación de conclusiones teóricas sobre la percepción del profesor, sino que aspira a una repercusión práctica más o menos inmediata. Una finalidad indirecta y última de una investigación de este tipo es mejorar la eficacia docente. De hecho, no han faltado investigaciones que han tratado de asociar la percepción docente a la eficacia, como por ejemplo las que describe Gage (1958). Si el profesor toma conciencia de los esquemas y filtros que utiliza para percibir a sus alumnos, estará más capacitado para corregir las posibles limitaciones de estas estructuras, si es que efectivamente distorsionan los atributos reales que caracterizan al alumno. La toma de conciencia comienza probablemente por la lectura de informes de investigación empírica que descubran la existencia, naturaleza y funcionalidad de los esquemas implícitos.

Referencias

- ALBRIGHT, L. KENNY, D. y MALLOY, T. (1988). Consensus in Personality judgments at zero acquaintance. *Journal of Personality and Social Psychology*, 55, 387-395.
- ANGUERA, M. T. (1983). *Manual de prácticas de observación*. México: Trillas.
- BERKOWITZ, L. (Ed.) (1988). *Advances in experimental social psychology* (Vol. 21). Nueva York: Academic Press.
- BORKENAU, P. (1986a). Systematic distortions in the recognition of trait information. En A. Angleitner, A. Furnham, y G. van Heck (Eds.). *Personality Psychology in Europe: Current trends and controversies* (pp. 177-191). Lisse: Swets y Zeitlinger.
- BORKENAU, P. (1986b). Toward and understanding of trait interrelations: Acts as instances for several traits. *Journal of Personality and Social Psychology*, 51, 371-381.
- BORKENAU, P. (1988). The multiple classification of acts and the Big Five Factors of personality. *Journal of Research in Personality*, 22, 337-352.
- BORKENAU, P. y OSTENDORF, F. (1987a). Retrospective estimates of act frequencies: How accurately do they reflect reality? *Journal of Personality and Social Psychology*, 52, 626-638.
- BORKENAU, P., y OSTENDORF, F. (1987b). Fact and fiction in Implicit Personality Theory. *Journal of Personality*, 55, 415-443.
- BROPHY, J. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, 5, 631-661.
- BRUNER, J., y TAGIURI, R. (1954). The perception of people. En G. Lindzey (Ed.). *Handbook of social psychology* (Vol. 1, pp. 634-654). Cambridge, MA.: Addison-Wesley.
- BUSS, D., y CRAIK, K. (1983). The act frequency approach to personality. *Psychological Review*, 90, 105-126.
- CADWELL, J., y JENKINS, J. (1985). Effects of the semantic similarity of items on student ratings of instructors. *Journal of Educational Psychology*, 77, 383-393.
- CANTOR, N., y MISCHEL, W. (1979a). Prototypes in person perception. En L. Berkowitz (Ed.). *Advances in Experimental Social Psychology* (Vol. 12, pp. 3-52). Nueva York: Academic Press.

- CANTOR, N., y MISCHEL, W. (1979b). Prototypicality and personality: Effects of free recall and personality impressions. *Journal of Research in Personality*, 13, 187-205.
- CHAN, D., y JACKSON, D. (1979). Implicit theory of psychopathology. *Multivariate Behavioral Research*, 14, 3-19.
- CHAN, D., y JACKSON, D. (1982). Individual Differences in the perception of judgment of psychopathology. *Multivariate Behavioral Research*, 17, 3-32.
- CONNOLLY, J., y DOYLE, A.-B. (1981). Assessment of social competence in preschoolers: Teachers versus peers. *Developmental Psychology*, 17, 454-462.
- CROCKER, J. (1981). Judgment of covariation by social perceivers. *Psychological Bulletin*, 90, 272-292.
- CRONBACH, L. (1955). Processes affecting scores on understanding of others and assumed similarity. *Psychological Bulletin*, 52, 177-193.
- DAROM, E., y RICH, Y. (1988). Sex differences in attitudes towards schools: Student self-reports and teacher perceptions. *British Journal of Educational Psychology*, 58, 350-355.
- DIGMAN, J., e INOUE, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology*, 50, 116-123.
- DIGMAN, J., y TAKEMOTO-CHOCK, N. (1981). Factors in the natural language of personality: Re-analysis and comparison of six major studies. *Multivariate Behavioral Research*, 16, 149-170.
- DORNBUSCH, S.; HASTORF, A.; RICHARDSON, S.; MUZZY, R., y VREELAND, R. (1965). The perceiver and the perceived: Their relative influence on categories of interpersonal perception. *Journal of Personality and Social Psychology*, 1, 434-440.
- EBBESON, E., y ALLEN, R. (1979). Cognitive processes in implicit personality trait inferences. *Journal of Personality and Social Psychology*, 37, 471-488.
- FRENCH, D., y WAAS, G. (1985). Teacher's ability to identify peer-rejected children: A comparison of sociometrics and teacher ratings. *Journal of School Psychology*, 23, 347-353.
- FURNHAM, A. (1984). Lay conceptions of neuroticism. *Personality and Individual Differences*, 5, 95-105.
- GAGE, N. (1985). Explorations in teachers perceptions of pupils. *Journal of Teacher Education*, 9, 97-101.
- GRANLEESE, J.; TURNER, I., y TREW, K. (1989). Teachers' and boys' and girls' perceptions of competence in the primary school: The importance of physical competence. *British Journal of Educational Psychology*, 59, 31-37.
- GUTTMANN, J., y BAR-TAL, D. (1982). Stereotypic perceptions of teachers. *American Educational Research Journal*, 19, 519-528.
- HAMILTON, D., y ROSE, T. (1980). Illusory correlation and the maintenance of stereotypic beliefs. *Journal of Personality and Social Psychology*, 39, 832-845.
- HAMPSON, S. (1982). Person memory: A semantic category model of personality traits. *British Journal of Psychology*, 73, 1-11.
- HAMPSON, S. (1983). Trait ascription and depth of acquaintance: The reference for traits in personality descriptions and its relation to target familiarity. *Journal of Research in Personality*, 17, 398-411.
- HASTIE, R., y KUMAR, P. (1979). Person memory: Personality traits as organizing principles in memory for behavior. *Journal of Personality and Social Psychology*, 37, 25-38.
- HUBER, G., y MANDL, H. (1979). Spiegeln Lehrerurteile über Schüler die implizite Persönlichkeitsstruktur der Beurteiler oder der Beurteilungsbögen? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 11, 218-231.
- ITSKOWITZ, R.; NAVON, R., y STRAUSS, H. (1988). Teachers' accuracy in evaluating students' self-image: effect of perceived closeness. *Journal of Educational Psychology*, 80, 337-341.
- JACKSON, D.; CHAM, D., y STRICKER, L. (1979). Implicit personality theory: Is it illusory? *Journal of Personality*, 47, 1-10.
- JACKSON, P. (1968). *Life in classrooms*. Nueva York: Holt, Rinehart, and Winston.
- KAHNEMAN, D.; SLOVIC, P., y TVERSKY, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases*. Nueva York: Cambridge University Press.
- KENEALY, P.; FRUDE, N., y SHAW, W. (1988). Influence of children's physical attractiveness on teacher expectations. *The Journal of Social Psychology*, 128, 373-383.
- KENRICK, D., y STRINGFIELD, D. (1980). Personality traits and the eye of the beholder: Crossing some traditional philosophical boundaries in the search for consistency in all of the people. *Psychological Review*, 87, 88-104.
- KIM, H., y BARON, R. (1988). Exercise and the illusory correlation: Does arousal heighten stereotypic processing? *Journal of Experimental Social Psychology*, 24, 366-380.
- KRUGLANSKI, A., y AJZEN, I. (1983). Bias and error in human judgment. *European Journal of Social Psychology*, 13, 1-44.
- LAMIELL, J.; FOSS, M., y CAVENEE, P. (1980). On the relationships between conceptual schemes and behavior reports: A closer look. *Journal of Personality*, 48, 54-73.

- LARSON, J. (1979). The limited utility of factor analytic techniques for the study of implicit theories in student ratings of teacher behavior. *American Educational Research Journal*, 16, 201-211.
- LAY, C., y JACKSON, D. (1969). Analysis of the generality of trait-inferential relationships. *Journal of Personality and Social Psychology*, 12, 12-21.
- LAY, C.; BURRON, B., y JACKSON, D. (1973). Base rates and information value in impression formation. *Journal of Personality and Social Psychology*, 28, 390-395.
- LEWICKI, P. (1982). Trait relationships: The nonconscious generalization of social experience. *Personality and Social Psychology Bulletin*, 8, 439-445.
- LOEVINGER, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635-694.
- MARKUS, H., y WURF, E. (1987). The dynamic self-concept: A socio-psychological perspective. En M. Rosenzweig y L. Porter (Eds.). *Annual Review of Psychology* (Vol. 38, pp. 299-337). Palo Alto, CA: Annual Reviews Inc.
- MAYER, J., y BOWER, G. (1986). Learning and memory for personality prototypes. *Journal of Personality and Social Psychology*, 51, 473-492.
- MILLER, D., y TURNBULL, W. (1986). Expectancies and interpersonal processes. En M. Rosenzweig y L. Porter (Eds.). *Annual Review of Psychology* (Vol. 37, pp. 233-256). Palo Alto, CA: Annual Reviews Inc.
- MIRELS, H. (1976). Implicit personality theory and inferential illusions. *Journal of Personality*, 44, 467-487.
- NISBETT, R., y ROSS, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.
- PARK, B. (1986). A method for studying the development of impressions of real people. *Journal of Personality and Social Psychology*, 51, 907-917.
- PASSINI, F., y NORMAN, W. (1966). A universal conceptions of personality structure? *Journal of Personality and Social Psychology*, 4, 44-49.
- REED, P., y JACKSON, D. (1975). Clinical judgment of psychopathology: A model for inferential accuracy. *Journal of Abnormal Psychology*, 84, 475-482.
- ROBERTS, A., y JENKINS, P. (1982). Teachers' perceptions of assertive and aggressive behavior at school: A discriminant analysis. *Psychological Reports*, 5, 827-832.
- ROSENBERG, S. (1989). A study of personality in literary autobiography: An analysis of Thomas Wolfe's «Look Homeward Angel». *Journal of Personality and Social Psychology*, 56, 416-430.
- ROSENBERG, S., y JONES, R. (1972). A method for investigating and representing a person's implicit theory of personality: Theodore Dreiser's view of people. *Journal of Personality and Social Psychology*, 22, 372-386.
- ROSENBERG, S., y SEDLAK, A. (1972). Structural representations of implicit personality theory. En L. Berkowitz (Ed.). *Advances in Experimental Social Psychology* (Vol. 6, pp. 235-297). Nueva York: Academic Press.
- ROSENBERG, S.; NELSON, C., y VIVEKANANTHAN, P. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, 9, 283-294.
- ROSENTHAL, R., y JACOBSON, L. (1968). *Pygmalion in the classroom*. Nueva York: Holt, Rinehart y Winston.)
- SCHNEIDER, D. (1973). Implicit personality theory: A review. *Psychological Bulletin*, 79, 294-305.
- SEMIN, G., y CHASSEIN, J. (1985). The relationship between higher order models and everyday conceptions of personality. *European Journal of Social Psychology*, 15, 1-15.
- SEMIN, G., y KRAHÉ, B. (1987). Lay conceptions of personality: Eliciting tiers of a Scientific conception of personality. *European Journal of Social Psychology*, 17, 199-209.
- SEMIN, G.; ROSCH, E., y CHASSEIN, J. (1981). A comparison of the common-sense and «scientific» conceptions of extraversion-introversion. *European Journal of Social Psychology*, 11, 77-86.
- SHROUT, P., y FLEISS, J. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.
- SHWEDER, R. (1975). How relevant is an individual difference theory of personality? *Journal of Personality*, 43, 455-485.
- SHWEDER, R. (1982). Fact and artifact in trait perception. The systematic distortion hypothesis. En B. Maher y W. Maher (Eds.) *Progress in Experimental Personality Research* (Vol. 11, pp. 65-100). Nueva York: Academic Press.
- SHWEDER, R., y D'ANDRADE, R. (1979). Accurate reflection or systematic distortion? A reply to Block, Weiss, and Thorne. *Journal of Personality and Social Psychology*, 37, 1075-1084.
- SHWEDER, R., y D'ANDRADE, R. (1980). The systematic distortion hypothesis. *New Directions for Methodology of Social and Behavioral Science*, 4, 37-58.

- SMITH, S., & KIHLSSTROM, J. (1987). When in a schema not a schema? The Big Five traits as cognitive structures. *Social Cognition, 5*, 26-57.
- SNYDER, M., & SWANN, W. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology, 36*, 1202-1212.
- STERNBERG, R.; CONWAY, B.; KETRON, J., & BERNSTEIN, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology, 41*, 37-55.
- STRICKER, L.; JACOBS, P., & KOGAN, N. (1974). Trait interrelations in implicit personality theories and questionnaires data. *Journal of Personality and Social Psychology, 30*, 198-207.
- SWEDE, S., & TETLOCK, P. (1986). Henry Kissinger's implicit theory of personality: A quantitative case study. *Journal of Personality, 54*, 617-646.
- TSOI, M., & NICHOLSON, J. (1982). A factor analytic study of teachers ratings of the personality of school children. *Personality and Individual Differences, 3*, 53-63.
- WATSON, D. (1989). Strangers' ratings of the Five Robust Personality Factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology, 57*, 120-128.
- WHITELY, S., & DOYLE, K. (1976). Implicit theories in student ratings. *American Educational Research Journal, 13*, 241-253.
- WINTER, L.; ULEMAN, J., & CUNNIFF, C. (1985). How automatic are social judgments? *Journal of Personality and Social Psychology, 49*, 904-917.