

La investigación en Psicología Diferencial Cognoscitiva: estrategias de segunda generación

R. COLOM MARAÑÓN Y M. DE JUAN ESPINOSA

Universidad Autónoma de Madrid



Resumen

En este artículo se exponen las estrategias de segunda generación que, dentro del ámbito de investigación de la Psicología Diferencial Cognoscitiva, se consideran más representativas, a saber: el análisis componencial, el análisis factorial confirmatorio, y el análisis de rasgo latente. Estas estrategias de investigación en Psicología Diferencial Cognoscitiva tienen, en contraste con las estrategias de primera generación, un dominio de aplicación más amplio. Tanto las estrategias de primera como las de segunda generación pueden ser utilizadas conjuntamente, siempre de acuerdo a las consideraciones teóricas de interés que trate de analizar el investigador.

Abstract

In this paper Second Generation Strategies are put forth. Within the research domain of Cognitive Differential Psychology, the analyses considered to be more representative are: componential analysis, confirmatory factor analysis, and latent trait analysis. These research strategies, in contrast with first generation strategies, have a broader domain of application. Strategies of both generations may be used jointly to the extent that their conjunction fits the relevant theoretical considerations that the researcher aims to analyse.

INTRODUCCIÓN

Hemos presentado en un artículo anterior (Juan Espinosa y Colom, 1989) un marco de interpretación de la investigación en psicología diferencial cognoscitiva. En ese artículo se establecía una clasificación entre las diversas estrategias de investigación en psicología diferencial cognoscitiva. Distinguíamos entre estrategias de primera generación, y estrategias de segunda generación. Entre las primeras teníamos la estrategia de correlatos, las estrategias heurísticas, y las estrategias de comparación de grupos. Las estrategias de correlatos se dividían en estrategias de correlatos biológicos y correlatos psicológicos. Las estrategias heurísticas se dividían en estrategia racional, estrategia de simulación por ordenador, y estrategia de simulación matemática. Por último, en las estrategias de comparación de grupos se distinguían las subestrategias de comparación de grupos con tareas de modelo implícito, las subestrategias de comparación de grupos con tareas de modelos explícitos, y la subestrategia de contraste mediante simulación diferencial. El rasgo distintivo de estas estrategias de primera generación era que estudiaban ámbitos concretos de las diferencias individuales, sin ánimo de elaborar una teoría general sobre dichas diferencias.

El término estrategias de segunda generación, por su parte, hace referencia a estrategias que se caracterizan por intentar aumentar el poder descriptivo y explicativo de las estrategias de primera generación, así como por tratar de incrementar el alcance de las potenciales aplicaciones derivadas de sus resultados. Es interesante, a este respecto, la siguiente afirmación de Hertzog (1985): «Hemos llegado a unos niveles tales de complejidad que la receta para entrenar adecuadamente a un psicólogo que pretenda estudiar la conducta inteligente podría ser la siguiente: un tercio de psicología experimental, un tercio de psicometría, y un tercio de tecnología estadística, sazonado con un poquito de psicología fisiológica e inteligencia artificial» (p. 59).

En el presente artículo, vamos a presentar tres de estas estrategias de segunda generación que consideramos representativas de entre las actualmente disponibles. Estas estrategias son: (a) el análisis componencial; (b) el análisis factorial confirmatorio, y (c) el análisis de rasgo latente. Estas tres estrategias de investigación tienen por objetivo explicar una serie de actividades cognoscitivas a partir de un mismo marco de referencia, tratando de que los resultados que se obtengan puedan ser encajados en dicho marco. Por ejemplo, la estrategia de investigación componencial se ha aplicado a campos como el razonamiento inductivo, el razonamiento deductivo, etc. El análisis de rasgo latente se ha aplicado al análisis diferencial de analogías verbales, al estudio de la aptitud verbal, de la aptitud numérica, etc. El análisis factorial confirmatorio se ha aplicado con cierto detenimiento al análisis de la lectura (Frederiksen, 1978).

Una cuestión que hay que hacer notar es que estas estrategias de segunda generación pueden servirse de las estrategias de primera generación; de hecho, esto ocurre así. Por ejemplo, el análisis componencial utiliza simulaciones matemáticas del comportamiento de los sujetos al razonar sobre problemas de inferencia transitiva. El análisis de rasgo latente utiliza las estrategias de comparación de grupos para detectar los procesos diferenciales de procesamiento. El análisis factorial confirmatorio utiliza la es-

trategia de correlatos. Asimismo, las estrategias de segunda generación pueden utilizarse recíprocamente: p.e., el análisis de rasgo latente, para detectar los componentes de procesamiento, utiliza las técnicas componenciales, aunque autores representativos (p.e., Embretson, 1985) afirme que los análisis componenciales *per se*, sirven de poco para modificar sustancialmente la estructura y sistematización de los tradicionales tests de inteligencia. Vamos a presentar, a continuación estas estrategias de investigación de segunda generación.

EL ANÁLISIS COMPONENCIAL

El análisis componencial es una estrategia de investigación que tiene por objetivo medir los procesos cognitivos subyacentes al rendimiento diferencial de los individuos en las tareas de los tests psicométricos de aptitudes (p.e., resolución de analogías, de silogismos lineales, de series incompletas, etc.). En líneas generales, esta estrategia de investigación consiste en seleccionar una tarea y descomponerla hasta llegar a los componentes más sencillos de procesamiento, hipotéticamente necesarios para la adecuada ejecución de la tarea completa. Estos componentes se secuencializan según un hipotético modelo, bien adoptado de otro ámbito de investigación, bien generado por el propio investigador. Siguiendo este procedimiento, Robert Sternberg (1979) encuentra que algunos componentes del procesamiento de la información correlacionan, lo que le permite postular, con cierta plausibilidad psicológica, la existencia de una serie de «metacomponentes», metacomponentes que tendrían un carácter similar al «funcionamiento ejecutivo» postulado por Belmont y Butterfield (1982), y que en última instancia harían alusión a los aspectos *metacognitivos* del funcionamiento mental.

Según John Carroll (1979), el análisis componencial implica situaciones experimentales que varían sistemáticamente las oportunidades del sujeto de comprometerse en diferentes tareas de procesamiento de la información. Por ejemplo, en una tarea de analogías $A : B :: C : (D)$, D representaría una serie de estímulos alternativos; sólo algunos muestran un valor de verdad. Sternberg (1977 a y b) presenta los estímulos en distintas secuencias temporales, de tal forma se puedan determinar los tiempos de procesamiento que un sujeto requiere en los diferentes estadios de resolución (p.e., latencias). Según Carroll (1979) el procedimiento del análisis componencial es una aplicación sofisticada del método de sustracción ideado por Donders (1868).

Para dar una visión panorámica, previa a la revisión sistemática del análisis componencial, vamos a servirnos de la descripción de Richard Snow (1979) sobre las fases que suele atravesar un análisis componencial:

1. Cualquier tarea se entiende como un conjunto de subtareas; de esta forma se obtiene una puntuación compuesta para cada individuo. El primer paso de un análisis componencial es *identificar* estas *subtareas*.
2. Después de identificar las subtareas, se realiza un experimento para obtener un intervalo de puntuaciones para cada subtarea, de acuerdo a los parámetros «latencia» y «error».
3. Los intervalos de puntuaciones se suponen *aditivos*.
4. A partir de los intervalos de puntuaciones, se estiman las puntua-

ciones del componente, usando uno o más modelos de procesamiento de la información de la tarea, para cada sujeto.

5. Se valida internamente el modelo de componentes, examinando las correlaciones entre las puntuaciones del componente, utilizando el concepto «validez de constructo», y realizando experimentos que tengan implicaciones para el modelo.

6. Se realiza un análisis de regresión para examinar la red de relaciones entre las puntuaciones del componente, del intervalo, y del factor.

7. Los pasos anteriores desembocan en un análisis intensivo de la tarea. Este tipo de análisis aplicado a un conjunto de tareas, daría lugar a un análisis extensivo de la tarea.

En resumen, el análisis componencial pretende identificar las operaciones mentales de los componentes subyacentes a una serie de tareas del procesamiento de la información relacionadas. También pretende descubrir la organización de esas operaciones componenciales. Este último objetivo se alcanzaría buscando la función de relación entre componentes y las constelaciones de aptitudes mentales. Los componentes son los elementos causales de estas relaciones, siendo *preceptivos* los componentes generales, y *opcionales* los componentes de grupo.

Por tanto, podemos comprobar que el análisis componencial pretende reunir las aportaciones «más interesantes» de los enfoques psicométricos y los enfoques cognitivos, para estudiar la inteligencia humana. Procedamos a continuación a una revisión detallada del análisis componencial, sirviéndonos de Sternberg (1985).

1. La primera etapa consiste en seleccionar o generar una teoría pertinente al fenómeno psicológico que pretendemos estudiar. Se trata de decidir cuál va a ser el objeto de nuestro análisis. Se requiere una teoría o modelo descriptivo y explicativo de ese objeto de análisis. En la selección o generación de una teoría, se deberían tener presentes los siguientes criterios: a) *completitud*: explicación de los procesos implicados en el objeto seleccionado; b) *especificidad*: descripción detallada del funcionamiento de los aspectos de la cognición; c) *generalidad*: aplicable a un amplio conjunto de problemas; d) *parsimonia*: explicación con el menor número de parámetros y asunciones de funcionamiento y e) *plausibilidad*: explica datos que prueban la teoría.

2. En la segunda etapa se seleccionan las tareas pertinentes para poder llevar a cabo el análisis. Las tareas seleccionadas deben satisfacer una serie de criterios:

El criterio de *cuantificabilidad* asegura la posibilidad de la asignación de numerales a objetos o eventos de acuerdo con una serie de reglas. Si se cuantifican aspectos de los protocolos, las cuantificaciones pueden ser variables dependientes pertinentes, en la medida en que satisfagan los otros criterios. El criterio de *fiabilidad* mide la variación de la puntuación verdadera en relación con la puntuación total. Debe calcularse para los ítems y para los sujetos. La *validez de constructo* asegura que la tarea ha sido elegida en consonancia con alguna teoría psicológica. La *validez empírica* asegura que la tarea sirve para el fin que se supone que sirve.

3. En tercer lugar, se descomponen la tarea en componentes de ejecución. Hay una serie de razones que justifican el aislamiento de los compo-

nentes del procesamiento de la información a partir de subtareas, y no de tareas completas, como por ejemplo, la posibilidad de aislar los componentes del Procesamiento de la Información a partir de subtareas que no pueden ser aisladas a partir de las tareas completas. Además, el uso de subtareas exige que el investigador especifique en qué subtarea se ejecuta cada componente del PI (tal y como veíamos en la estrategia de simulación por ordenador; véase Juan Espinosa y Colom, 1989). De esta forma, se requiere una especificación de las relaciones entre la estructura de la tarea y los componentes asociados. El uso de subtareas incrementa el número de datos que han de explicarse, lo que nos permite evitar un posible ajuste espúreo entre el modelo y los datos, ajuste que puede producirse cuando el número de parámetros a estimarse es demasiado grande en relación al número de datos que han de predecirse. El uso de subtareas da lugar a estimaciones de componentes libres de ejecución para una serie de intervalos de procesamiento anidados.

Entre los modos de descomposición de tareas que se pueden utilizar en un análisis componencial, se encuentran los siguientes (Sternberg, 1985).

El *método de las tareas completas* (Guyote y Sternberg, 1981) es el más sencillo, y no exige el supuesto de aditividad. No obstante, en muchas de las tareas se confundirán los componentes del PI. Tales confusiones pueden llevar a graves consecuencias (cf. Sternberg, 1977 b). Es un método adecuado para items en los cuales no hay confusión de los componentes.

El *método de pre-indicadores* (Sternberg, 1977 a y b; Sternberg, 1980 b; Sternberg y Gardner, 1983) permite desentrañar los componentes, comparar los modelos, y aumentar el número de datos de que se puede disponer en la elaboración de modelos. Este método exige que el investigador especifique en qué intervalo de procesamiento tiene lugar cada operación mental, y proporciona puntuaciones del rendimiento en una serie de intervalos del procesamiento. No obstante, exige el supuesto de aditividad, la prueba individual, y no es adecuado para niños pequeños debido a su complejidad. El método de los preindicadores sigue la siguiente lógica. En primer lugar, se deben formar puntuaciones de intervalo a partir de la descomposición de la tarea completa en una serie de subtareas relacionadas. En segundo lugar, se considera que cada puntuación de intervalo es una puntuación en alguna de las subtareas, y mide la ejecución en algún subconjunto de los componentes del PI requeridos para la tarea completa. Por tanto, cada subtarea requiere sucesivamente menos procesamiento con lo que se reducirá el tiempo y la dificultad del procesamiento.

El *método de las tareas parciales* (Sternberg, 1980 a y b; Sternberg y Turner, 1981) comparte todas las ventajas del método de los pre-indicadores, y sólo la desventaja del supuesto de aditividad. Por lo tanto, se debe utilizar preferentemente cuando se pueda usar cualquiera de los dos métodos. No obstante, se puede necesitar alguna prueba piloto para determinar qué método es más probable que produzca la aditividad a través de las condiciones. Algunas tareas pueden descomponerse por cualquiera de los dos métodos, pero no todas. En consecuencia, la decisión de qué método usar sólo puede efectuarse tras una cuidadosa consideración de las demandas de la tarea y de su descomponibilidad. La lógica de este método consiste en presentar items completos. Los indicadores no se descomponen en dos partes,

a diferencia del método de los preindicadores, sino que la descomposición se efectúa en ensayos unitarios.

El método del «*stem-splitting*» (Sternberg y Nigro, 1980) podría utilizarse para pruebas colectivas, es utilizable con niños pequeños, y parece crear un cierto interés añadido en los sujetos. No obstante, el método no ha sido verificado adecuadamente; su generalidad a otros problemas que no sean analogías no se han demostrado todavía, y parece más adecuado qué los precedentes para generar estrategias especiales inaplicables a las tareas estándar (completas). Supone items que requieren el mismo número de componentes del PI, y un número diferente de ejecuciones de los diversos componentes; combina características de los métodos anteriores.

El método de los cuadernillos variados sistemáticamente (Sternberg y Rifkin, 1979) es práctico con niños muy pequeños, y se puede pasar colectivamente. No obstante, no es posible obtener una medida pura del tiempo empleado sólo en items contestados correcta o incorrectamente, y no es particularmente adecuado para desentrañar los componentes. La unidad de presentación es el cuadernillo, y se les da a los sujetos un intervalo temporal determinado para completar tantos items como puedan dentro de un determinado cuadernillo. La clave del método es que todos los items del cuadernillo deben ser homogéneos respecto a la teoría o teorías que están siendo contrastadas. Los items son heterogéneos en los distintos cuadernillos.

4. En cuarto lugar, el análisis componencial cuantifica el modelo, por ejemplo, mediante diagramas de flujo del procesamiento de la información (véase la Figura 2).

La cuantificación supone generalmente utilizar la regresión múltiple para predecir los parámetros de la variable dependiente, a partir de la/s variable/s independiente/s. La variable dependiente será el tiempo de respuesta y el porcentaje de error. Las variables independientes serán el número de veces que se ejecuta cada uno de los componentes del procesamiento de la información. Los parámetros de latencia representan las duraciones de los diversos componentes. El Tiempo de Reacción se hipotetiza igual a la suma de las cantidades de tiempo empleadas en cada operación. Se formula la hipótesis de que la proporción de errores es igual a la suma de las dificultades encontradas al ejecutar cada una de las operaciones. En los modelos de tiempo de respuesta, todas las operaciones deben contribuir significativamente a la latencia de solución, ya que, por definición, cada ejecución de una operación consume alguna cantidad de tiempo. En la investigación del razonamiento analógico de Sternberg (1977 a y b) por ejemplo, el modelamiento matemático se hizo por medio de la regresión lineal múltiple y los parámetros del modelo se estimaron como coeficientes de regresión no estandarizados.

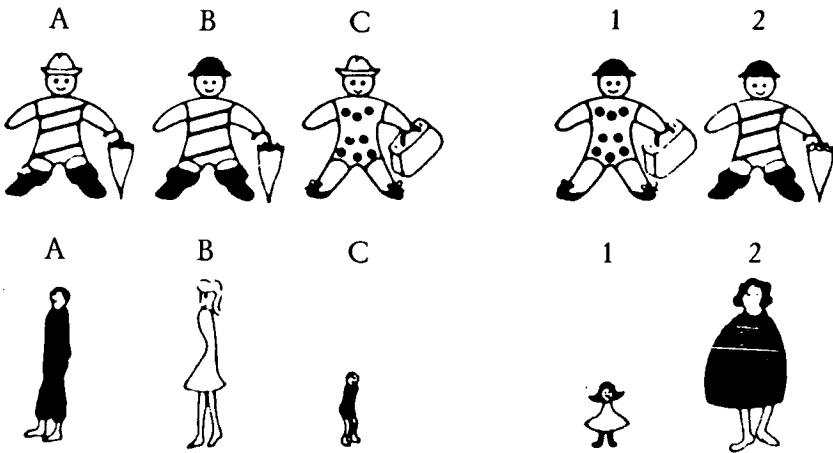
5. En quinto lugar tendremos que contrastar el modelo, lo que es tanto como decir que tendremos que analizar la validez interna de nuestro modelo. Las pruebas para realizar la validación interna de un modelo componencial irán desde el escrutinio del R^2 del modelo, el RMSD del modelo, la regresión del modelo, etc.

Las tres etapas generales del análisis componencial son la contrastación del modelo (p.e., análisis de la validez externa), es decir, en qué medida es útil para proporcionar una fuente adicional de verificación del modelo; pro-

porciona una prueba de la generalidad del modelo propuesto. La reformulación del modelo supone realizar una validación cruzada del mismo. Una vez se completa la reformulación, el modelo ha de contrastarse con los nuevos datos. Por último, hay que tratar de generalizar el modelo. El proceso de generalización es necesario a fin de establecer la prioridad de la teoría del PI, más que del análisis de la tarea *per se*.

Veamos una ejemplificación de cómo se desarrolla un análisis componencial. Concretamente vamos a presentar La Teoría Componencial del Razonamiento Analógico (versión de Sternberg y Rifkin, 1979, remodelada a partir de la teoría presentada por Sternberg en 1977 a y b).

FIGURA 1



Ejemplos de Analogías utilizadas en la investigación de Sternberg y Rifkin (1979). Arriba se muestran analogías esquemáticas y abajo analogías integrales (véase el texto).

Teóricamente, el razonamiento analógico implica seis procesos, cuatro de los cuales son obligatorios, y dos son opcionales (a diferencia de la teoría presentada en 1977, en la que sólo el proceso «justificación» era opcional). Los procesos son los siguientes (veáanse los ejemplos de la Figura 1): 1) **CODIFICACIÓN** (proceso obligatorio): el sujeto codifica un término de la analogía y almacena en la MCP (a) atributos posiblemente relevantes del término, y (b) un valor correspondiente a cada atributo almacenado. En el ejemplo de la Figura 1, cada término contiene cuatro atributos y dos posibles valores para cada atributo: altura (alto o bajo), peso (gordo o delgado), color del vestido (blanco o negro), sexo (hombre o mujer). 2) **INFERENCIA** (proceso obligatorio): el sujeto descubre la relación entre los términos A y B de la analogía y almacena la relación entre los mismos en la MCP. En el ejemplo de la Figura 1, esta relación se puede representar mediante la expresión altura (0), peso (0), color del vestido (blanco o negro), sexo (hombre o mujer). La designación de conjunto nulo (0) se utiliza para indicar una transformación nula. 3) «**MAPPING**» proceso opcional): el sujeto conecta el dominio (primera parte) de la analogía con el rango (segunda parte) de la analogía, descubriendo la relación entre los términos A y C. En el ejemplo, esta relación se puede representar mediante la expresión altura (alto o bajo), peso (0), color del vestido (0), Sexo (0). 4)

APLICACIÓN (proceso obligatorio): el sujeto aplica desde C a todas las opciones de respuesta una relación análoga a la relación inferida en el dominio de la analogía. En las dos analogías del ejemplo de la Figura 1, esta relación se puede construir sólo para C y 1. Los contenidos de la relación son idénticos a los de la relación inferida. 5) **JUSTIFICACIÓN** (proceso opcional): en algunas analogías, ninguna opción de respuesta permite una relación C-opción que se corresponda exactamente con la relación extraída de A-B. En tales analogías, el sujeto debe justificar la superioridad de alguna de las opciones. El sujeto lo hace determinando qué opción de respuesta permite una relación que difiere sólo ligeramente de la relación extraída del dominio. Por último, se selecciona esta opción, de respuesta. 6) **RESPUESTA**: el sujeto da una solución —1 ó 2—.

Los Modelos de Combinación de los Componentes son de dos clases. La primera clase de modelos se aplica a las analogías con atributos integrales (parte inferior de la Figura 1), y la segunda clase de modelos se aplica a las analogías con atributos separables o analogías esquemáticas (parte superior de la Figura 1). En un estímulo con atributos separables, es posible tener un valor nulo para cualquiera de los atributos, y aún así preservar la integridad del estímulo. Un valor nulo en cualquiera de los atributos significa simplemente que el atributo no existe. La segunda analogía de la Figura 1 es un ejemplo de analogía con atributos integrales. Los atributos integrales y separables están sujetos a mecanismos psicológicos diferentes. Sternberg y Rifkin (1979) hablan de *procedimientos* y *estrategias*. Un procedimiento es un modelo de procesamiento de la información, ni opcional ni consciente. Una estrategia es un modelo de procesamiento de la información consciente y opcional.

Presentan cuatro Modelos Procedimentales. Se asume que los procesos se ejecutan serialmente (véase el Diagrama de Flujo de la Figura 2). Además, se asume que la codificación de los términos de la analogía se realiza en modo exhaustivo de procesamiento. Los sujetos almacenan en la MLP todos los atributos y los valores correspondientes que ellos creen pueden ser relevantes para resolver la analogía. Cuando así lo exige la fase de resolución, estos atributos y valores se recuperan en MCP. Los modelos difieren en la secuenciación de los procesos, y en el número de veces que se realizan los procesos de inferencia, mapping, y aplicación. Cada uno de estos procesos se puede ejecutar de manera exhaustiva o auto-terminada. Si un proceso se ejecuta en modo *exhaustivo*, se realizan comparaciones entre todos los valores de los atributos codificados para determinado par de estímulos. Si un proceso se ejecuta en modo *auto-terminado*, sólo se realiza un subconjunto de las comparaciones posibles.

Según el Modelo I, los procesos de inferencia, *mapping*, y aplicación se realizan en modo exhaustivo. Los sujetos infieren todas las posibles relaciones entre los atributos codificados de los dos primeros términos de la analogía; a continuación los sujetos proyectan (*maps*) todas las posibles relaciones entre los atributos codificados entre el primer y tercer término de la analogía. Finalmente, el sujeto aplica todas las posibles relaciones entre el tercer término y cada una de las opciones de respuesta. En el ejemplo de la Figura 1, el sujeto infiere en primer lugar que el color del vestido y el sexo (pero nada más) cambia de A a B; a continuación el sujeto proyecta (*maps*) la relación entre A y C, reconociendo que sólo cambia la altura; fi-

FIGURA 2

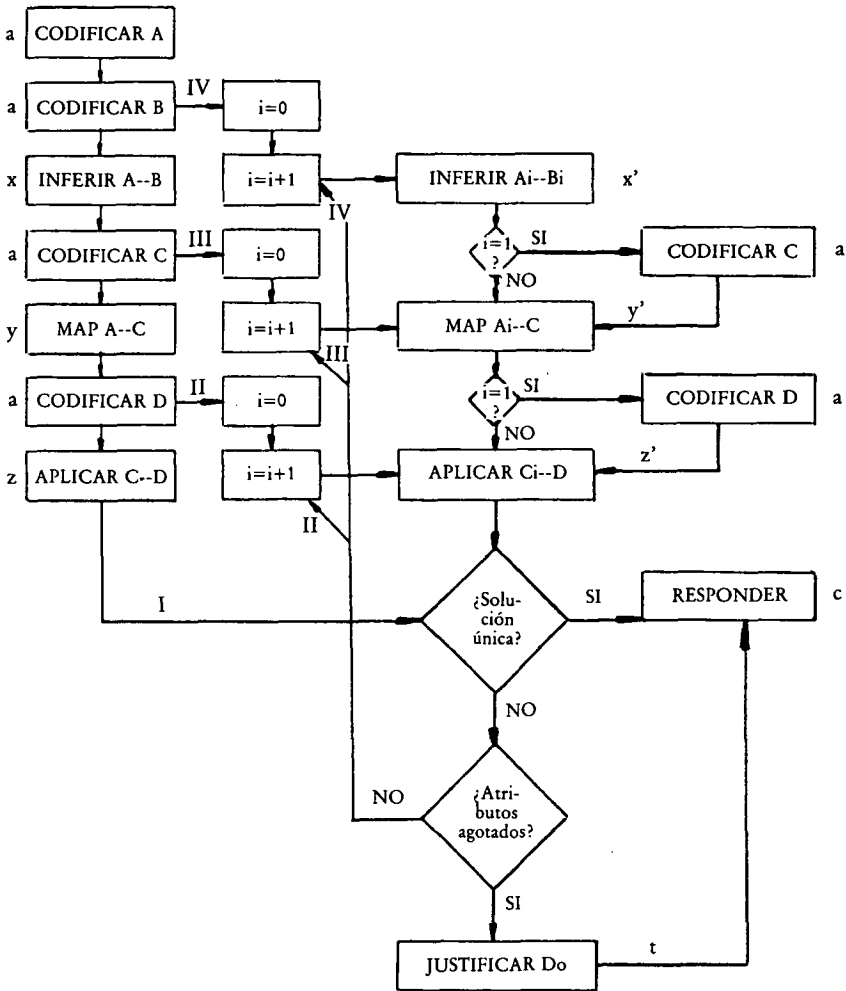


Diagrama de Procesamiento de los Modelos presentados en el texto (tomado de Sternberg y Rifkin, 1979).

nalmente, el sujeto aplica la regla inferida de C, a cada una de las opciones de respuesta, reconociendo que cambian los mismos valores de los atributos (y permanecen constantes) de C a 1, al igual que de A a B.

De acuerdo al Modelo II, los procesos de inferencia y *mapping* se realizan en modo exhaustivo y la aplicación se realiza en modo auto-terminado. El sujeto infiere en modo exhaustivo todas las relaciones entre los atributos codificados de A y B, y proyecta (*maps*) en modo exhaustivo todas las relaciones entre los atributos codificados de A y B. El sujeto sólo aplica tantos valores de los atributos como sea necesario para elegir la respuesta. El orden de selección de los atributos es aleatorio. Supongamos que el primer atributo seleccionado es el sexo. Los sujetos aplican «sexo» desde el término C a cada una de las opciones de respuesta, pero se encuentra con que es incapaz de elegir una opción de acuerdo a este atributo, puesto que

las dos opciones son mujeres. A continuación, supongamos que los sujetos seleccionan «altura». La opción de respuesta debe ser «bajo» y sólo la primera opción corresponde a «bajo». Por tanto, la respuesta es 1. Hay que observar que el sujeto resuelve la analogía de acuerdo a sólo dos aplicaciones, y no se extiende en analizar los cuatro atributos, como exigiría el Modelo I.

Por lo que se refiere al Modelo III, el proceso de inferencia se realiza en modo exhaustivo, y los procesos de *mapping* y aplicación se realizan en modo auto-terminado. El sujeto infiere en modo exhaustivo todas las relaciones entre los atributos codificados de A y B. A partir de este momento, el sujeto cambia a modo auto-terminado. Proyecta un valor del atributo de A a C, y aplica el valor del atributo correspondiente de C a D. Si el atributo elegido es suficiente para distinguir entre las opciones de respuesta correcta e incorrectas, el sujeto responderá. En otro caso, el sujeto proyecta y aplica otro atributo, tratando de seleccionar una sola respuesta. En el ejemplo de la Figura 1, supongamos que el sujeto trata de proyectar en modo auto-terminado el atributo sexo. A y C son del mismo sexo, pero C y D son de diferente sexo. Este conocimiento es insuficiente para elegir entre las opciones de respuesta, debido a que ambas opciones son del mismo sexo. De esta forma, el sujeto debe proyectar otro atributo, p.e., la altura. Posteriormente, el sujeto aplica «altura» de C a D, encontrando que sólo la opción 1 tiene el valor necesario (p.e., «bajo»). Ahora, el sujeto puede responder.

Por último, según el Modelo IV, los procesos de inferencia, *mapping*, y aplicación se realizan en modo auto-terminado. En primer lugar, el sujeto infiere un valor del atributo de A a B, y proyecta el valor del atributo correspondiente de A a C; finalmente, aplica el valor del atributo de C a cada una de las opciones de respuesta. Si el sujeto es capaz de distinguir la respuesta correcta de la incorrecta, responderá. En otro caso, el sujeto deberá inferir, proyectar y aplicar otro valor, y continuará hasta que encuentre el valor que distinga la opción correcta de la incorrecta. En el ejemplo, supongamos que el sujeto comienza con el atributo «sexo». En primer lugar, el sujeto infiere los cambios en el sexo de A a B; a continuación proyecta el valor constante para el sexo de A a C. En la aplicación, el sexo cambia de C a la opción correcta de respuesta, pero ambas opciones satisfacen el criterio; por tanto, debe repetirse el bucle. Supongamos que el sujeto selecciona «altura» como el siguiente atributo a contrastar. Este atributo puede proporcionar la base para distinguir entre las opciones de respuesta. Por tanto, el sujeto podrá responder.

En resumen, los Modelos I, II, III y IV, difieren en qué procesos (inferencia, proyección y aplicación) se realizan en modo exhaustivo, y cuáles se realizan en modo auto-terminado. En el Modelo I, las tres operaciones se realizan en modo exhaustivo; en el Modelo II, la inferencia y la proyección se realizan en modo exhaustivo, y la aplicación se realizan en modo auto-terminado; en el Modelo III, sólo la inferencia se realiza en modo exhaustivo; en el Modelo IV, todos los procesos se realizan en modo auto-terminado.

Por otro lado, los Modelos de Estrategia se etiquetan con el sufijo «M», para indicar que son modificaciones de los modelos originales (estas modificaciones no aparecen en la exposición de 1977). Estos modelos difieren

de los modelos procedimentales en dos aspectos fundamentales. Ambas diferencias se derivan de la separabilidad de los atributos estímulares procesados por estos modelos. Esta separabilidad facilita la manipulación de atributos únicos aislados. En primer lugar, no se necesita el proceso de proyección. Se asume que los sujetos pasan directamente de la inferencia a la aplicación. Por ejemplo, en el ejemplo de la Figura 1, el sujeto infiere en primer lugar que el color del sombrero cambia (de blanco a negro). El sujeto aplica entonces este cambio de C a cada una de las opciones de respuesta, buscando la opción en la que cambie el color del sombrero. La proyección de A a C se suprime. En segundo lugar, se asume que la codificación pasa a modo auto-terminado inmediatamente antes del primer proceso de comparación de atributos auto-terminado (sea inferencia o aplicación). Cuando la codificación pasa a modo auto-terminado, los sujetos codifican atributos únicos de cada término de la analogía sólo hasta donde sea necesario para realizar los procesos de comparación de atributos. Sólo se codifican los atributos absolutamente necesarios para resolver la analogía. Si, por ejemplo, un sujeto debe inferir un cambio en el color del sombrero, sólo se codificará el color del sombrero antes de la inferencia. Sólo se codificarán otros atributos adicionales si el color del sombrero no sirve para distinguir entre las opciones de respuesta.

El Modelo IM difiere del Modelo I sólo en la ausencia de proyección. Se asume que los sujetos infieren y aplican los valores en modo exhaustivo. En el Modelo II-IIIM, la inferencia se realiza en modo exhaustivo, y la aplicación se realiza en modo auto-terminado. La codificación auto-terminada comienza con el término C. Las modificaciones de los modelos II y III se combinan, debido a que el proceso de proyección no se realiza en modo exhaustivo o auto-terminado. En el Modelo IVM, la inferencia y la aplicación se realizan en modo auto-terminado, la codificación auto-terminada comienza con el procesamiento del término A de la analogía.

Esta Teoría Componencial del Razonamiento Analógico ha sido puesta recientemente a prueba con sujetos con retraso mental por McConaghy y Kirby (1987). Su objetivo es averiguar si el entrenamiento mejora el rendimiento de sujetos con retraso mental en tareas de razonamiento analógico, siguiendo la teoría componencial de Sternberg. Estos autores encuentran que este entrenamiento produce un mayor lapso temporal de codificación, y un menor número de errores.

Según McConaghy y Kirby (*op. cit.*) un examen de los resultados obtenidos por Sternberg y Rifkin (1979) muestra que todos los modelos (I, II, III y IV) proporcionan ajustes similares a los datos para todos los grupos de edad estudiados por los dos últimos autores. Por su parte, McConaghy y Kirby (1986) encuentran un resultado similar utilizando sujetos por encima y por debajo del promedio, y sujetos retrasados. Esta similitud de ajuste a los modelos «parece una dificultad seria del método componencial» (McConaghy y Kirby, 1987, p. 22).

EL ANÁLISIS FACTORIAL CONFIRMATORIO

El análisis factorial tradicional es por naturaleza *exploratorio*, trata de descubrir la naturaleza de un conjunto de constructos subyacentes median-

te un examen inductivo de la estructura factorial de un conjunto de medidas observadas. Por el contrario, el análisis factorial confirmatorio es un *contraste de hipótesis* que utiliza las concepciones teóricas sobre la naturaleza de los constructos y sus relaciones con las variables observadas para especificar (*a priori*) un modelo que prediga la estructura factorial de un conjunto de medidas. Las concepciones deben basarse en la investigación empírica previa, así como en conjuntos de nociones teóricas, o una combinación de ambos aspectos. La idea es que registran datos intentando desconfirmar o falsar un determinado modelo o conjunto de modelos generados a partir de una teoría, mostrando que la matriz de correlaciones obtenida empíricamente no es consistente con el modelo propuesto. La distinción exploratorio/confirmatorio supone métodos de utilización del análisis factorial para realizar inferencias a partir de los datos empíricos. La distinción restringido/no-restringido, por su parte, se refiere al tipo de algoritmo factor-analítico utilizado para estimar los parámetros del modelo del análisis factorial (Hertzog, 1985).

La utilidad derivada de distinguir el análisis factorial confirmatorio del análisis restringido proviene del hecho de que los métodos de análisis factorial restringido de Joreskog (1974) se utilizan frecuentemente en estudios que sólo pueden caracterizarse como exploratorios por naturaleza. Inversamente, algunos de los pasos del análisis factorial confirmatorio (p.e., selección cuidadosa de las variables empíricas para obtener una representación del dominio factorial de interés válida en cuanto al contenido, dado un conjunto de hipótesis sobre la naturaleza de la estructura factorial subyacente) se han discutido ampliamente en la literatura antes del advenimiento de los métodos restringidos de análisis factorial (p.e., Thurstone, 1947). La razón por la que se hace énfasis en los métodos de análisis factorial restringido es que parecen estar elaborados para las investigaciones del análisis factorial confirmatorio, y disponen de un conjunto potente de instrumentos para la inferencia estadística, no disponibles en la mayor parte de los métodos no restringidos tradicionales. De todas formas, algunos de los aspectos más importantes del análisis factorial confirmatorio tienen poco o nada que ver con los métodos de análisis factorial restringido *per se*, y la utilización de *Cofamm*, *Lisrel*, o programas similares no es suficiente para producir un análisis factorial confirmatorio interesante.

Modelo de Factor Común: desarrollo general

Tanto el análisis factorial restringido como el no-restringido operan de acuerdo al mismo modelo de factor común. Supongamos un modelo que relacione las variables observadas «p» y los factores comunes no-observados. La representación usual es:

$$\begin{aligned} Y_1 &= \lambda_{1,1} \eta_1 + \lambda_{3,2} \eta_2 + \dots + \lambda_{1,m} \eta_m + \varepsilon_1 \\ Y_2 &= \lambda_{2,1} \eta_1 + \lambda_{2,2} \eta_2 + \dots + \lambda_{2,m} \eta_m + \varepsilon_2 \\ Y_3 &= \lambda_{p,1} \eta_1 + \lambda_{p,2} \eta_2 + \dots + \lambda_{p,m} \eta_m + \varepsilon_p \end{aligned}$$

que representa cada variable observada como una combinación lineal de los factores, denotados η_1, η_2 , etc., y un componente residual, ε_p . El patrón de coeficientes del factor (cargas), $\lambda_{p,m}$ son regresiones de las variables obser-

vadas en los factores. En notación matricial, la ecuación anterior se presenta como:

$$Y = \Lambda\eta + \varepsilon$$

donde se incluyen las variables observadas p en Y , los factores m en η , los componentes únicos en ε , y en Λ es una matriz de coeficientes de regresión $p \times m$. La ecuación anterior puede utilizarse para derivar la expectativa de la matriz de covarianza de Y , Σ , suponiendo que η y ε no correlacionan, que es la representación usual de los residuos de la regresión. El modelo de factor común representa Σ como:

$$\Sigma = \Lambda\psi\Lambda' + \Theta_\varepsilon$$

donde Λ es como antes, ψ es la matriz de covarianza de los factores, y Θ_ε es la matriz de covarianza de los componentes únicos. En otras palabras, los elementos diagonales de Θ_ε son las varianzas únicas (varianza no explicada por los factores comunes) mientras que los elementos no diagonales son covarianzas entre los componentes únicos. Igualmente, los elementos diagonales de ψ son las varianzas de los factores, mientras que los elementos no diagonales son las covarianzas entre los factores.

El modelo de la ecuación anterior se ha extendido a un modelo para el análisis simultáneo de varios grupos. Simplemente se añade un exponente g para indicar la especificación factorial en el grupo « g -ésimo»:

$$\Sigma(g) = \Lambda(g)\psi(g)' + \Theta_\varepsilon(g)$$

Por tanto, el modelo anterior será un caso especial para un solo grupo.

Soluciones factoriales restringidas y no-restringidas

La mayor diferencia entre el análisis factorial restringido y el no-restringido radica en la forma en que estos métodos estiman los parámetros de la última ecuación. Dado que los factores η , y los residuos ε no son observados, hay un número infinito de soluciones en términos de los parámetros que podrían reproducir la matriz de covarianza Σ . La aproximación no-restringida no restringe en la parte factorial común $\Lambda\psi\Lambda'$. Por el contrario, especifica Θ_ε como matriz diagonal de varianzas únicas (p.e., no se asumen correlaciones entre los componentes únicos) y (temporalmente) específica como matriz identidad. Entonces, dada una especificación del número de factores, un conjunto de métodos familiares se utiliza para estimar la matriz de patrones factoriales no-restringidas. Esta matriz de patrones factoriales también es no-singular y puede estar sujeta a cualquier transformación arbitraria (rotación) para llegar a la solución «final».

La aproximación restringida es diferente. En primer lugar, todos los parámetros en las tres matrices paramétricas se incluyen en el modelo. Se procede especificando un modelo para la estructura factorial de las variables observadas en Y :

- a) Hipotetizando el número de factores *a priori*.
- b) Restringiendo los valores de un número suficiente de parámetros en las tres matrices de parámetros para permitir una estimación de los parámetros no conocidos.

Hay dos tipos de restricciones; los parámetros están:

- a) Fijados previamente a algún valor específico (por hipótesis).
- b) Algún conjunto de parámetros desconocidos se constriñen a un valor similar.

Por ejemplo, si hipotetizamos que una variable no está determinada (al menos en parte) por un determinado factor, podemos fijar la regresión de esa variable, en el factor, en un valor cero. Si hipotetizamos que dos variables tienen exactamente la misma magnitud de relación (distinta de cero) respecto a determinado factor, podemos representar esa hipótesis haciendo iguales las cargas factoriales asociadas. Suponiendo que se haya proporcionado un conjunto suficiente de restricciones respecto a determinado modelo para proporcionar una única solución, se pueden estimar los factores restantes.

La aproximación restringida se denomina de tal forma debido a que ubica restricciones en el espacio factorial común; es necesario ubicar un número suficiente de restricciones en los parámetros en Λ y ψ para identificar singularmente las desconocidas. De esta forma, la validez de las estimaciones paramétricas del modelo restringido depende de la validez de las restricciones ubicadas en el modelo.

¿Se produce que la carga de la variable 1 en el factor 2 se fija en cero, dada la hipótesis sobre las propiedades del factor? Una de las implicaciones de la aproximación restringida es que las variables observadas se deben seleccionar cuidadosamente para el análisis factorial, de forma que se pueda especificar previamente un modelo factorial significativo. A continuación se presentan las características destacadas de los modelos restringidos de análisis factorial, puesto que tiene mucho que ver con el proceso del análisis factorial confirmatorio.

Estadios del análisis factorial confirmatorio

Sin ninguna duda la fase más crítica del análisis factorial confirmatorio es la primera: una formulación clara de las hipótesis a contrastar y su impacto sobre el diseño factorial subsecuente. La racionalidad del análisis factorial confirmatorio se basa en el supuesto de que no se puede descubrir inductivamente la «verdadera» estructura de la inteligencia (o cualquier otra cosa) mediante el análisis exploratorio de pruebas seleccionadas de una forma un tanto arbitraria. Por el contrario, lo que se necesita es especificar previamente una hipótesis sobre la estructura factorial de un conjunto de pruebas seleccionadas cuidadosamente, dada una teoría factorial sobre la manera de determinar las diferencias individuales en la prueba. La lógica de la inferencia en el análisis confirmatorio es, esencialmente, la *falsación*. Si nos aseguramos de que los resultados del análisis factorial confirmatorio se pueden utilizar *de facto* para la contrastación de hipótesis, debemos asegurar una conexión válida entre las hipótesis respecto a la naturaleza de los constructos estudiados (p.e., la estructura factorial de algún dominio de variables observadas) y el diseño del estudio factorial. Aspectos a considerar:

1. Selección de un subconjunto de variables medidas a partir del dominio de variables observadas que miden el factor.
2. Consideración de las propiedades de las variables observadas, tal y

como definen un determinado espacio factorial. Esta consideración incluye un muestreo suficiente para permitir:

a) La representación de las covarianzas significativas entre las pruebas en el espacio factorial común mediante el muestreo suficiente de las variables observadas.

b) Identificación de las diferencias entre los dominios asegurando los cálculos suficientes en el «hiperplano de referencia».

3. Decisiones cuidadosas respecto de la población objetivo definida por el interés de la investigación.

4. Muestreo de forma que se asegure la viabilidad estadística del análisis confirmatorio, así como la validez externa de los resultados.

Por ejemplo, Widiger, Knudson y Rorer (1980) contrastan la hipótesis de que el factor «estilo cognitivo de independencia de campo» (Vernon, 1972) no es en realidad nada más que la reinterpretación del factor perceptual de Flexibilidad de Clausura, medido por el ETS de Ekstrom, French, Harman y Dermen (1976) utilizando el análisis factorial confirmatorio. Las pruebas estadísticas utilizadas no son nada más que una prueba de las relaciones empíricas entre las variables seleccionadas como representantes de los dos factores. Antes de registrar cualquier dato, hemos de averiguar si la Independencia de Campo se puede medir tal y como se hace tradicionalmente (p.e., *Rod and Frame Test*). La validez del diseño dependerá de las ramificaciones de la selección de variables. Si no hay defectos previos, después de registrar los datos, se determina el análisis. Una de las mayores ventajas del análisis factorial confirmatorio es que obliga a tener en cuenta una serie de consideraciones previas, antes de registrar cualquier dato, debido a que esas consideraciones son necesarias para determinar el modelo a contrastar, y por tanto, el diseño a ejecutar.

En conclusión, el término confirmatorio se refiere a la consistencia empírica, y no tanto a la validez de la teoría que produce el modelo. En principio, el análisis factorial restringido permite la evaluación de la consistencia empírica para muestras independientes, y proporciona los medios para la contrastación estadística de las hipótesis previamente formuladas, de acuerdo a determinado modelo teórico de base.

EL ANÁLISIS DE RASGO LATENTE

Los métodos de rasgo latente se pueden aplicar a cualquier teoría que sea susceptible de matematización. Dos de los objetivos de estos métodos son:

a) Analizar los componentes que supuestamente explican las diferencias individuales del rendimiento en los tests, explicando los mecanismos teóricos subyacentes.

b) Proporcionar parámetros para evaluar la contribución de los constructos teóricos específicos a la dificultad de la tarea.

Según Embretson (1985) aunque el análisis de componentes cognitivos (p.e., el análisis componencial de Sternberg) es una estrategia de investigación interesante, los tests de inteligencia no se han visto prácticamente modificados por los resultados de la aplicación de esa estrategia de investiga-

ción. Una hipótesis que podría explicar este fenómeno, es que los tests psicométricos se basan en principios claramente diferentes a los fundamentos de una teoría cognitiva. También podría deberse a que quizá son tan útiles en la consecución de sus fines, que no necesitan modificarse, aun a pesar de sus defectos teóricos.

El test tradicional no mide un único constructo teórico, y la heterogeneidad de constructos resultante no se detecta necesariamente acudiendo al criterio de consistencia interna propio de la psicometría tradicional. La estrategia de componentes cognitivos, por su parte, permite analizar con más profundidad este fenómeno: si los componentes se representan de forma similar en todos los ítems de un determinado test, o si los componentes correlacionan en la población, dicho test tendrá cierta consistencia interna. Un problema que se puede presentar es que subconjuntos del mismo tipo de ítem pueden depender desigualmente de varios componentes. En este sentido, Whitely (1981) muestra que los ítems se pueden seleccionar para medir diferentes capacidades de procesamiento (p.e., algunos ítems pueden depender de la inferencia de una regla, y otros pueden depender de la valoración cuidadosa de las alternativas de respuesta). El mismo tipo de ítem en tests diferentes, o la inclusión de ítems nuevos en un determinado *pool* de ítems, puede producir una medida diferente de la misma capacidad analizada.

En los Tests Adaptativos Computerizados (CAT) que construye el equipo de Susan Embretson, los tests no tienen un contenido fijo. La capacidad se mide presentando a los individuos ítems referidos a su nivel de capacidad. Es decir, *los tests tienden a individualizarse*. En este sentido, la teoría cognitiva puede ser útil para guiar la especificación de los ítems, y para calibrar los ítems en la elaboración de tests individualizados.

Para Susan Embretson (1985), dos de los aspectos más importantes para explicar la naturaleza de las diferencias individuales en un test son:

1. Los componentes, las estrategias y la base de conocimientos implicados en la resolución de un determinado ítem; p.e., representación de constructo.
2. Cuáles sirven a la validez predictiva del test; p.e., amplitud nomotética.

Embretson (1985) se plantea que si es posible separar las representaciones de constructo y la amplitud nomotética, también será posible incorporar los constructos y métodos de la investigación experimental en el proceso de validación de constructo. Evidentemente, esto puede ser una buena estrategia para unificar las dos disciplinas de la psicología científica (Cronbach, 1957, 1975).

Un método que pretenda evaluar la representación de constructo en un test, necesita alcanzar cuatro criterios:

1. Las hipótesis tienen que ser pertinentes respecto de, o relevantes para, los constructos implicados en la ejecución de la tarea.
2. Se deben especificar las relaciones entre las características de la tarea y las diferencias individuales.
3. Se deben estimar los parámetros que van a representar las diferencias individuales en los constructos psicológicos considerados.

4. Se deben estimar los parámetros, para representar la cualidad de los ítems en relación a los constructos considerados.

Estos cuatro criterios se alcanzan con los modelos de rasgo latente (véase la justificación en Embretson, 1983).

Por otra parte, los *modelos de ecuaciones estructurales* (Joreskog, 1974) tienen las propiedades adecuadas para investigar la amplitud nomotética (son útiles para evaluar los cuatro aspectos de la investigación de la amplitud nomotética reseñados más arriba). Los requisitos que deben cumplir los constructos considerados, en la investigación de la amplitud nomotética, son:

1. Los *pesos relativos* de los constructos, en la explicación de las diferencias individuales detectadas en las puntuaciones obtenidas en los tests.
2. La *bondad* con la que explican los constructos, la amplitud de las relaciones entre las puntuaciones de los tests y otras variables.

Veamos con más detalle las peculiaridades de la representación de constructo, y la amplitud nomotética.

La representación de constructo y la amplitud nomotética

Las fases que plantea Embretson (1985) para analizar la representación de constructo son las siguientes:

1. En primer lugar, se deben operativizar los constructos teóricos en un diseño experimental, de forma que se puedan poner a prueba las hipótesis sobre la tarea que hemos formalizado.
2. Se deben generar, a partir de la/s teoría/s de base, predicciones sobre el rendimiento en la tarea.
3. Se deben poner a prueba las predicciones de teorías alternativas a la que hayamos adoptado o generado.

Se deben revisar las teorías y la investigación previas para generar hipótesis sobre las variables procesuales que supuestamente subyacen al rendimiento. Posteriormente, se deben diseñar tareas que representen los constructos identificados. Se deben obtener predictores de la dificultad de la tarea, para el modelo matemático a elaborar.

Los métodos del diseño de tareas para descomponer variables teóricas, varían:

1. El contenido de la tarea (p.e., la complejidad estimular, o las características semánticas).
2. Las condiciones de presentación de la tarea (p.e., porcentaje, claridad de los estímulos, o instrucciones).
3. Completitud de la tarea (p.e., presentación de tareas parciales, o de pre/indicadores).

Por ejemplo, la revisión de Whitely (1973) de la investigación factorial sobre analogías verbales, encuentra resultados muy inconsistentes según diferentes estudios. Embretson (1985) muestra que ítems del mismo tipo pueden no necesariamente requerir las mismas capacidades, de forma que las inter/correlaciones entre los ítems, y sus correlaciones con otros tipos de ítems, podrían variar.

Veamos un ejemplo de descomposición de tareas, tal y como se utiliza en las investigaciones de Susan Embretson y sus colaboradores.

Embretson, Schneider y Roth (1985) han aplicado los métodos de descomposición de tareas a analogías verbales, para tratar de analizar los componentes cognitivos que subyacen a la ejecución ostensible. La revisión de la literatura indica que al menos tres estrategias están implicadas en la resolución de analogías verbales:

A. Estrategia orientada según la regla (véase Whitely y Barnes, 1979). De acuerdo a esta estrategia, el sujeto infiere la solución ideal a partir de las relaciones de la cadena del ítem, y valora las alternativas de respuesta de acuerdo a esa inferencia. La *construcción de la regla* y la *valoración de la respuesta* representan los dos componentes de esta estrategia.

B. Estrategia de asociación. Muchos ítems de los tests de aptitudes se pueden resolver seleccionando la opción de respuesta más asociada con el término C de la cadena, sin tener en cuenta la analogía definida por el par relacionado (p.e., dominio de la analogía).

C. Estrategia de eliminación de la respuesta. Representa el procesamiento realizado para resolver ítems de elección múltiple. Los estudios previos (p.e., Whitely y Barnes, 1979; Pelegrino y Glaser, 1980; Embretson y Curtright, 1982) indican que el procesamiento está influido por el número y tipo de alternativas de respuesta. Embretson y Curtright (1982) elaboraron una teoría del procesamiento de las alternativas de respuesta en dos estadios: (1) eliminación de las alternativas de respuesta que no pertenecen al dominio de la respuesta ideal; (2) si aún queda más de una alternativa, aparece un segundo estadio en el que se aplica una regla más discriminante.

Los parámetros del modelo incluyen capacidades para cada sujeto en cada componente, y para cada ítem en cada componente. Estos parámetros permiten la predicción de la efectividad de la estrategia para cada sujeto en cada componente. De esta forma, la predicción del encuentro de una persona con un ítem, también incluye el diagnóstico sobre la estrategia que con mayor probabilidad utilizará una persona al resolver un ítem.

Por lo que se refiere a la amplitud nomotética, la intención es averiguar la utilidad de un test como una medida de las diferencias individuales. La amplitud nomotética se mide por el patrón, magnitud, y frecuencia de las correlaciones del test con otras medidas de diferencias individuales, tales como otras características psicológicas, o el rendimiento en algunas tareas criterio.

Si se identifican múltiples componentes en el análisis de la representación de constructo, se deben responder varias preguntas sobre las diferencias individuales:

1. Cuál es la contribución de los componentes de procesamiento a las diferencias individuales.
2. Examinar la contribución de los constructos considerados, a la validez predictiva del test.
3. Determinar si los componentes de procesamiento tienen validez diferencial, a la hora de predecir los criterios pre-establecidos. Si es así, la medida de los componentes individuales incrementará la validez predictiva del test.

4. Generalizabilidad entre tareas, para establecer la significación de las medidas componenciales. Las diferencias individuales detectadas en el mismo componente, pero en diferentes tipos de ítems, estarán muy correlacionados. La investigación de la amplitud nomotética puede dar cuenta de esta temática estudiando las correlaciones de las diferencias individuales en los componentes medidos por diferentes tareas. Esto proporciona una operación convergente a los métodos experimentales, para establecer la generalizabilidad de los componentes entre tareas.

Los modelos de rasgo latente

Los modelos de rasgo latente, básicamente, conectan la teoría psicológica y las medidas de diferencias individuales, combinando modelos matemáticos y modelos psicométricos. En este apartado vamos a limitarnos a exponer brevemente las líneas argumentales básicas de los modelos de rasgo latente para proporcionar una guía general que sirva de marco. Comenzaremos con el modelo básico (p.e., el modelo de Rasch) para que se puedan comprender los ulteriores desarrollos y sofisticaciones a que ha dado lugar. Como hemos comentado más arriba, estos modelos sirven a cualquier teoría psicológica que sea susceptible de matematización, por lo que tiene un rango de aplicación muy extenso.

El modelo de Rasch

Fundamentalmente, es un modelo de la respuesta al ítem; predice el encuentro de la persona j con el ítem i mediante dos parámetros, a saber: la capacidad de la persona, θ_j , y la dificultad del ítem, ϵ_i . La probabilidad de que el encuentro sea exitoso, $P(X_{ij} = 1)$, se ajusta al siguiente modelo:

$$P(X_{ij} = 1) = \frac{e^{(\theta_j - \epsilon_i)}}{1 + e^{(\theta_j - \epsilon_i)}}$$

Es decir, la capacidad de la persona, θ_j , y la dificultad del ítem, ϵ_i , se combinan aditivamente para determinar el potencial de respuesta. El modelo postula una relación compensatoria entre la capacidad y la dificultad de la tarea. Es decir, una persona de baja capacidad puede tener el mismo potencial de respuesta que una persona de alta capacidad, si a la persona de baja capacidad se le presentan ítems sencillos.

Los parámetros del modelo de Rasch se pueden estimar por varios métodos (p.e., máxima verosimilitud condicional). Los parámetros del ítem y de la persona son similares en magnitud a las puntuaciones Z .

Una característica importante de los modelos de rasgo latente es la separabilidad de los parámetros de la persona y el ítem. Esto permite matematizar las puntuaciones en el test para cualquier subconjunto de ítems, bajo el supuesto de que se ha ponderado su nivel de dificultad. La matematización de los tests es especialmente crucial para los Tests Adaptativos Computerizados (CAT). Esto no se puede hacer en la teoría clásica de tests, puesto que se centra en las varianzas y covarianzas, y no en la respuesta a los ítems.

El modelo logístico lineal de rasgo latente (véanse los trabajos de Scheiblechner, 1972; y Fischer, 1973).

El modelo logístico lineal difiere del modelo de Rasch en que la dificultad del ítem se modela de acuerdo a las propiedades estimulares del ítem.

Por tanto, se necesitan datos ajenos al encuentro de la persona y el ítem para poder identificar los parámetros. Específicamente, los ítems deben puntuarse en K factores de complejidad para determinar la dificultad del ítem. Por ejemplo:

$$\varepsilon_i = \sum_k p_k q_{ik} + d$$

donde q_{ik} = La puntuación del ítem i en el factor k.

p_k = El peso del factor k en la dificultad del ítem i.

d = Una constante de normalización.

Los factores del ítem pueden representar cualquier sistema teórico que puntúe o clasifique los ítems.

Este modelo conecta la dificultad del ítem con variables teóricas y conserva las propiedades del modelo de Rasch como modelo psicométrico. Una ventaja de este modelo es su capacidad de contrastar hipótesis sobre los modelos matemáticos. Es un modelo explicativo, puesto que trata de explicar la dificultad del ítem según un pequeño conjunto de parámetros (a diferencia del modelo de Rasch que estima un parámetro para cada ítem).

El modelo logístico lineal es un modelo multicomponencial puesto que se postulan múltiples estímulos o factores componentes subyacentes a la dificultad del ítem. No obstante, es un modelo *unidimensional* puesto que sólo postula un parámetro de capacidad para una persona. Al igual que en el modelo de Rasch, la capacidad y la dificultad del ítem tienen una relación compensatoria. No obstante, si varía el rendimiento de una persona en los factores de predicción (como suele ocurrir generalmente), se necesita un modelo multidimensional.

El modelo multicomponencial de rasgo latente

A diferencia del modelo logístico lineal de rasgo latente, el modelo multicomponencial postula que tanto los ítems como las personas varían en el potencial de respuesta. Se requiere información independiente respecto de los componentes necesarios para resolver un ítem. Este modelo necesita datos sobre los componentes, para poder identificar los parámetros. Estos datos se pueden obtener presentando subtareas que correspondan a cada componente.

A continuación se presentan subtareas que representan diferentes componentes de la resolución de analogías verbales. Las analogías tienen el formato «A:B::C:_____» (A es a B lo que C es a). Vamos a presentar una analogía verbal estándar, y dos componentes globales, Construcción de la Regla y Valoración de la Respuesta (cf. Whitely, 1981).

Ítem completo

Gato:Tigre::Perro:_____

a) León; b) Lobo; c) Ladrido; d) Perrito; e) Caballo

Construcción de la regla

Gato:Tigre::Perro:_____

Regla

Valoración de la respuesta

Gato:Tigre::Perro:_____

a) León; b) Lobo; c) Ladrillo; d) Perro; e) Caballo

Regla: Un canino salvaje

Subtarea de asociación

Perro

a) León; b) Lobo; c) Ladrillo; d) Perro; e) Caballo

El modelo multicomponencial especifica cómo se relacionan los componentes, con la probabilidad de que un determinado ítem estándar sea resuelto por una determinada persona. El modelo multicomponencial puede variar el modelo matemático, para que se corresponda con teorías alternativas del rendimiento en la tarea.

Los *modelos multicomponenciales* son una familia de modelos que especifican:

- a) Modelos matemáticos para relacionar el rendimiento para el ítem completo, y el rendimiento en determinado componente.
- b) Modelos psicométricos que relacionan el rendimiento en determinado componente, con la capacidad de la persona y la dificultad del ítem en ese componente.

Implicaciones de la matematización:

1. Los ítems y las personas tienen asignados parámetros en cada componente. Lo que mide el ítem completo, depende del patrón compuesto en relación a la dificultad de los componentes. Si un ítem es muy fácil, respecto de un componente, las diferencias individuales en la resolución del ítem total no dependerá de ese componente. Por tanto, este componente tendrá un efecto techo. La selección de ítems de este tipo no dará cuenta de las diferencias individuales en el componente. Por tanto, los ítems deben seleccionarse para medir las diferencias individuales en componentes específicos.

2. La causa de la dificultad de un ítem para determinada persona se debe diagnosticar explícitamente. Si una persona se puntúa en varios componentes, se deben proporcionar datos exactos en un determinado ítem. Se reproduce en el modelo la probabilidad de que la persona j resuelva cada componente k en un determinado ítem i . Las personas con diferentes patrones componenciales tendrán diferentes patrones de dificultad en el ítem completo.

3. Los componentes no son compensatorios. Se requieren todos los componentes para resolver el ítem. Si una persona tiene una baja probabilidad de lograr el resultado correcto en un determinado componente, las probabilidades altas de resolver otros componentes no compensan el componente fallido.

Las hipótesis sobre los modelos matemáticos se pueden poner a prueba mediante la bondad de ajuste de los tests. Los parámetros del ítem son útiles para diseñar un test, debido a que los ítems se pueden seleccionar para medir componentes específicos. Si se seleccionan ítems muy fáciles en componentes irrelevantes, las diferencias individuales en la resolución del ítem, reflejarán sólo los componentes que son difíciles (Whitely, 1981; Embretson, 1985).

El modelo general de rasgo latente (véase Embretson, 1984)

Combina los dos modelos anteriores, de la siguiente manera:

$$P(X_{ijT} = 1) = \Pi_k P^*(X_{ijk} = 1)$$

donde

$$P^*(X_{ijk} = 1) = e^{(\theta_{jk} - \sum_m q_{imk} \eta_{mk} + d_k)} / 1 + e^{(\theta_{jk} - \sum_m q_{imk} \eta_{mk} + d_k)}$$

En el modelo general, a cada respuesta componencial subyace un modelo de la complejidad del estímulo. Por tanto, los resultados componenciales se combinan de acuerdo al modelo matemático multicomponencial. De esta forma, la explicación se produce a dos niveles:

- a) El resultado en la tarea completa se conecta a los resultados en los componentes.
- b) El resultado en el componente se conecta a los factores de complejidad de las subtareas que definen el componente.

Retomemos el ejemplo de la analogía anterior. El componente construcción de la regla consiste en varias operaciones más elementales, como la codificación del significado de los términos, la inferencia de la relación entre Gato y Tigre, y la contextualización de la relación respecto a Perro. Los items pueden variar en la dificultad de estas operaciones, y en las características estímulares. Por ejemplo, la dificultad de proyección puede verse influida por el número de relaciones A:B educidas, debido a que la probabilidad de seleccionar una relación equivocada puede incrementarse.

Whitely y Curtright (1980) obtuvieron una serie de índices de varios factores estímulares para un *pool* de items de analogías. Se obtuvieron índices para cada ítem en tres factores postulados como determinantes de la dificultad de construcción de la regla. Estos factores son:

1. La facilidad de inferir la relación A:B, q_{i11} .
2. La amplitud relacional de las inferencias A:B (p.e., el número de relaciones educidas de A:B, q_{i21}).
3. La facilitación de la inferencia por el contexto que supone el término C, q_{i31} . Se obtuvo el siguiente modelo de construcción de la regla según la dificultad del ítem en el modelo general de rasgo latente:

$$\varepsilon_{ik} = -1.84q_{i11} + .06q_{i21} - 1.05q_{i31} + .68$$

Multiplicando los índices para cada ítem por estos pesos se obtiene un modelo de su dificultad respecto a la construcción de la regla. Por ejemplo, en la analogía anterior, se obtuvieron los índices .47, 2.17, .04, para la educación de la inferencia, la amplitud relacional, y la contextualización, respectivamente, según las normas de Whitely y Curtright (*op. cit.*). Estos valores se aplicarían a los pesos de la ecuación anterior.

La estimación de los parámetros del modelo general de rasgo latente se pueden obtener mediante el programa *Multicomp* (Whitely y Nieh, 1981).

Según Embretson (1985) los items deberían seleccionarse para desarrollar un test que:

1. Mida sólo las diferencias individuales en la ejecución de una estrategia, y contenga puntuaciones que representen un pequeño conjunto de habilidades cognitivas.

2. Mida varias estrategias, y contenga puntuaciones que representen una combinación variada de estrategias, dependiendo de cuál aplique el individuo.

3. Mida la combinación más válida de estrategias para predecir criterios importantes de validez.

Por tanto, el constructor de tests puede elegir entre incrementar el significado cognitivo del test, o incrementar la validez predictiva.

DISCUSIÓN Y CONCLUSIONES

Hemos visto una serie de estrategias de investigación que, tal y como dijimos en la introducción, tienen por objetivo analizar una serie de actividades cognitivas para tratar de elaborar teorías generales sobre las diferencias individuales, o el rendimiento diferencial de los individuos en tareas cognitivas de diversa índole. Estas estrategias de investigación son las que en Juan Espinosa y Colom (1989), se denominan estrategias de segunda generación.

Si revisamos el panorama presentado en éste, y en el anterior artículo citado, veremos una serie de ofertas para la investigación de las diferencias individuales. El investigador puede utilizar cualesquiera de ellas (sean de primera o de segunda generación), siempre y cuando haya realizado un análisis previo de las necesidades teóricas que pretende cubrir, de forma que la investigación venga *teóricamente guiada*, tratando de subsanar de esta forma el viejo problema de la psicología diferencial: la tendencia tecnológica que señala, por ejemplo, Hunt, Frost y Lunneborg (1973). Tendencia que no es en sí perniciosa, pero que debería correr pareja a desarrollos teóricos, tal y como se indica en Juan Espinosa y Colom (1989).

Además de tener en consideración el análisis de las necesidades particulares de determinado investigador a la hora de seleccionar la estrategia de investigación, se tendría que tener presente un mínimo marco de referencia teórico sobre las diferencias individuales. Un análisis de un posible marco de este tipo se presenta en Juan Espinosa y Colom (en prensa). De esta forma se evitaría la excesiva dispersión de resultados en la investigación de las diferencias humanas.

Referencias

- BELMONT, J. M., y BUTTERFIELD, E. C. (1982): To secure transfer of training, instruct self-management skills. En D. K. Detterman y R. J. Sternberg (comps.): *How and how much can intelligence be increased?* Norwood, NJ: Ablex Publishing Corporation.
- CARROLL, J. B. (1979): How shall we study individual differences in cognitive abilities? Methodological and theoretical perspectives. En R. J. Sternberg y D. K. Detterman (comps.): *Human intelligence: perspectives on its theory and measurement*. Norwood, N. J.: Ablex Publishing Corporation.
- CRONBACH, L. J. (1957): The two disciplines of scientific psychology. *American Psychologist*, 12, 671-684.
- CRONBACH, L. J. (1975): Beyond the two disciplines of scientific psychology. *American Psychologist*, 30, 116-127.
- DONDERS, F. C. (1868): La vitesse des actes psychiques. *Archives Néerlandaises*.
- EKSTROM, R. B.; FRENCH, J. W.; HARMAN, H. H., y DERMAN, D. (1976): *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- EMBRETSON, S. (1983): Construct validity: construct representation versus nomothetic span. *Psychological Bulletin*, 93, 179/197.
- EMBRETSON, S. (1984): A general latent trait model for response processes. *Psychometrika*, 49, 175-186.
- EMBRETSON, S. (1985): Intelligence and its measurement: extending contemporary theory to existing tests. En C. R. Reynolds y V. L. Wilson (comps.): *Methodological and statistical advances in the studies of individual differences*. Nueva York: Plenum Press.
- EMBRETSON, S., y CURTRIGHT, C. (1982): Problem structure and response format in solving verbal analogies. Informe técnico No. NIE-82-2 para el National Institute of Education, Lawrence, Kansas: Universidad de Kansas.
- EMBRETSON, S.; SCHNEIDER, L., y ROTH, D. (1985): Multiple processing strategies and the construct validity of verbal reasoning test. *Journal of Educational Measurement*.
- FISCHER, G. H. (1973): The linear logistic model as an instrument in educational research. *Acta Psychologica*, 37, 359-374.
- FREDERIKSEN, J. R. (1978): A chronometric study of component skills in reading. Cambridge, MA: Bolt Beranek and Newman, Inc. (Informe n.º 3757, Informe técnico n.º 2, ONR Contract N00014-76-C-0461, NR 154 386).
- GUYOTE, M. J., y STERNBERG, R. J. (1981): A transitive-chain theory of syllogistic reasoning. *Cognitive Psychology*, 13, 461-525.
- HERTZOG, C. (1985): Applications of confirmatory factor analysis to the study of intelligence. En D. K. Detterman (comp.): *Current topics in human intelligence, vol. 1: Research methodology*. Nueva Jersey: Ablex Publishing Corporation.
- HUNT, E.; FROST, N., y LUNNEBERG, C. (1973): Individual differences in cognition: a new approach to intelligence. En C. Bower (comp.): *Advances in learning and motivation, vol. 7*. Nueva York: Academic Press.
- JORESKOG, K. G. (1974): Analyzing psychological data by structural analysis of covariance matrices. En D. H. Krantz, R. C. Atkinson, R. D. Luce y P. Suppes (comps.): *Contemporary developments in mathematical psychology, vol. 2*. San Francisco: CA, Freeman.
- JUAN ESPINOSA, M., y COLOM, R. (1989): La Investigación en Psicología Diferencial Cognoscitiva: estrategias de primera generación. *Estudios de Psicología*.
- JUAN ESPINOSA, M., y COLOM, R. (en prensa): Aproximaciones a una Psicología Diferencial Cognoscitiva. *Cognitiva*.
- McCONAGHY, J., y KIRBY, N. H. (1986): Analogical reasoning and ability level: An examination of R. J. Sternberg's componential method. *Intelligence*.
- McCONAGHY, J., y KIRBY, N. H. (1987): Using the componential method to train mentally retarded individual to solve analogies. *American Journal of Mental Deficiency*, 92 (1), 12-23.
- PELLEGRINO, J., y GLASER, R. (1980): Components of inductive reasoning. En R. Snow, P. A. Federico y W. E. Montague (comps.) *Aptitude, learning and instruction: cognitive process analyses of aptitude, vol. 1*. Hillsdale, NJ: LEA.
- RASCH, G. (1960): *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- SCHIEBLECHNER, H. (1972): Learning and solving of complex mental problems. *Zeitschrift Für Experimentelle Angewandte Psychologie*, 19, 476-506.
- SNOW, R. (1979): Theory and method for research on aptitude Processes. En R. J. Sternberg y D. K. Detterman (comps.): *Human intelligence: Perspectives on its theory and measurement*. Nueva Jersey, Ablex Publishing Corporation.
- STERNBERG, R. J. (1977 a): *Intelligence, information processing, and analogical reasoning: the componential analysis of human abilities*. Nueva Jersey: LEA.
- STERNBERG, R. J. (1977 b): *Component processes in analogical reasoning*. *Psychological Review*, 84, 353-378.

- STERNBERG, R. J. (1979): The nature of mental abilities. *American Psychologist*, 34, 214-130.
- STERNBERG, R. J. (1980 a): Representation and Poces in linear syllogistic reasoning. *Journal of Experimental Psychology: General*, 109, 119-159.
- STERNBERG, R. J. (1980 b): Sketch of a componential subtheory of human intelligence. *Behavioral and Brain Sciences*, 3, 573-614.
- STERNBERG, R. J. (1985): Componential analysis: a recipe. En D. K. Detterman (comp.): *Current topics in human intelligence, vol. 1: Research methodology*. Nueva Jersey: Ablex Publishing Corporation.
- STERNBERG, R. J., y GARDNER, M. K. (1983): Unities in inductive reasoning. *Journal of Experimental Psychology: General*, 112, 80-116.
- STERNBERG, R. J., y NIGRO, G. (1980): Developmental patterns in the solution of verbal analogies. *Child Development*, 51, 27-38.
- STERNBERG, R. J., y RIFKIN, B. (1979): The development of analogical reasoning processes. *Journal of Experimental Child Psychology*, 27, 195-232.
- STERNBERG, R. J., y TURNER, M. E. (1981): Components of syllogistic reasoning. *Acta Psychologica*, 47, 245-265.
- THURSTONE, L. L., (1947): Multiple factor analysis. Chicago, Il: University of Chicago Press.
- VERNON, P. E. (1972): The distinctiveness of field independence. *Journal of Personality*, 40, 366-391.
- WHITELY, S. (1973): *Types of relationships in reasoning by analogy*. Tesis doctoral inédita. Minnesota: Universidad de Minnesota.
- WHITELY, S. (1981): Measuring aptitude processes with multicomponent latent trait models. *Journal of Educational Measurement*, 18, 67-84.
- WHITELY, S., y BARNES, G. M. (1979): The implications of processing events sequences for theories of analogical reasoning. *Memory and Cognition*, 7, 323-331.
- WHITELY, S., y CURTRIGHT, C. A. (1980): Performance and stimulus complexity norms for verbal analogy test items. Informe Técnico NIE-80-4 para el National Institute of Education, Lawrence, Kansas: Universidad de Kansas.
- WHITELY, S., y NIEH, K. (1981): Program Multicomp. Manuscrito inédito. Lawrence, Kansas: Universidad de Kansas.
- WIDIGER, T. A.; KNUDSON, R. M., y R ORER, L. G. (1980): Convergent and discriminant validity of measures of cognitive styles and abilities. *Journal of Personality and Social Psychology*, 39, 116-129.