

# MODELOS MENTALES<sup>1</sup>

(Mental models)

**Marco Antonio Moreira**  
Instituto de Física, UFRGS  
Caixa Postal 15051, Campus  
91501-970 Porto Alegre, RS, Brasil

## RESUMEN

El tema de los modelos mentales es enfocado particularmente a la luz de la teoría de Johnson-Laird. La visión de otros autores es también presentada, pero el énfasis del texto está en el enfoque de Johnson-Laird, quien propone los modelos mentales como una tercera vía para el debate imágenes x proposiciones. En esta perspectiva, se discute la naturaleza, el contenido y la tipología de los modelos mentales, así como la cuestión de la conciencia y la computabilidad. Además de eso, se focaliza también la metodología de la investigación en modelos mentales y son dados ejemplos de investigaciones. La intención del texto es, principalmente, la de servir como introducción al asunto de los modelos mentales con vistas a la investigación en enseñanza de las ciencias.

## ABSTRACT

The subject of mental models is presented particularly in the light of Johnson-Laird's theory. Views from different authors are also presented, but the emphasis lies in Johnson-Laird's approach, proposing mental models as a third path for the images x propositions debate. In this perspective, the nature, content, and typology of mental models are discussed, as well as the issue of consciousness and computability. In addition, the methodology of research on mental models is also discussed and examples of research studies are provided. Essentially, the aim of the text is to provide an introduction to the mental models topic, aiming at science education research.

*The mind must be more complicated than any theory of it: however complex the theory, a device that invented it must be still more complex (P. Johnson-Laird).*

## Introducción

El estudio de la mente no es cosa nueva. A pesar de que, como dice Johnson-Laird (1983), la mente siempre será más complicada que cualquier teoría respecto a ella, hace siglos que el hombre inventa teorías sobre cómo funciona su propia mente. Tanto es así que hoy en día se acepta la existencia de una Ciencia Cognitiva, o sea, un estudio interdisciplinario de la mente abarcando psicología cognitiva, lingüística, inteligencia artificial, neurociencia, filosofía y antropología cognitiva. Según Thagard (1996, p. 10), la hipótesis central de la ciencia cognitiva es que el funcionamiento de la mente es mejor entendido en términos de representaciones mentales y procedimientos computacionales que operan sobre dichas representaciones. Es decir, se supone que

---

<sup>1</sup> Programa Internacional de Doctorado en Enseñanza de las Ciencias. Universidad de Burgos, España; Universidade Federal do Rio Grande do Sul, Brasil. *Texto de Apoyo n° 8*. Originalmente presentado en el Encuentro sobre Teoría e Investigación en Enseñanza de Ciencias - Lenguaje, Cultura y Cognición, Facultad de Educación de la UFMG, Belo Horizonte, 5 a 7 de marzo de 1997. Publicado en portugués en *Investigações em Ensino de Ciências*, Porto Alegre, v. 1, n. 3, p. 193-232. Traducción de M<sup>a</sup> Luz Rodríguez Palmero. Revisado en 1999.

*la mente es representacional y computacional*. Para Thagard (ibid.) no hay duda de que la *comprensión representacional-computacional* de la mente es el más exitoso enfoque desarrollado hasta hoy para la mente, tanto teórica como experimentalmente.

Naturalmente, si el conocimiento en la mente consiste en *representaciones mentales* no debe sorprender que los científicos cognitivos hayan propuesto varios tipos de representaciones mentales como, por ejemplo, las reglas, los “scripts” y las imágenes discutidas en el texto número 7 (Greca, 1999) de esta serie. En este octavo texto, sin embargo, el foco estará en un nuevo constructo representacional: *los modelos mentales*, principalmente según la óptica de Philip Johnson-Laird (op. cit.) con el fin de ayudar en la enseñanza y en la investigación en enseñanza de las ciencias a la luz de este referente.

## **Representaciones mentales**

Las representaciones internas, o *representaciones mentales*, son maneras de “re-presentar” internamente el mundo externo. Las personas no captan el mundo exterior directamente; construyen representaciones mentales (es decir, internas) del mismo.

En principio, se puede distinguir entre representaciones mentales *analógicas* y *proposicionales*. La imagen visual es el ejemplo típico de representación analógica, pero hay otras como las auditivas, las olfativas, las táctiles.

Las *representaciones analógicas* son no-discretas (no-individuales), concretas (representan entidades específicas del mundo exterior), organizadas por reglas débiles de combinación y específicas de la modalidad a través de la que se encontró originalmente la información (Eisenck y Keane, 1994, p. 184).

Las *representaciones proposicionales* son discretas (individuales), abstractas, organizadas según reglas rígidas y captan el contenido ideacional de la mente independientemente de la modalidad original en la que se encontró la información, en cualquier lengua y a través de cualquiera de los sentidos. (ibid.).

Estas representaciones son “tipo-lenguaje”, pero se trata de un lenguaje que no tiene que ver con la lengua ni con la modalidad de percepción; es un lenguaje de la mente que podríamos llamar “mentales”. Las representaciones proposicionales no son frases en un lenguaje determinado. Son entidades individuales y abstractas formuladas en un lenguaje propio de la mente.

Hay psicólogos cognitivos para los que la cognición debe analizarse exclusivamente en términos de representaciones proposicionales, o sea, no hay necesidad de suponer que las imágenes son un tipo especial, separado, de representación mental. Para éstos, “los proposicionalistas”, las imágenes pueden reducirse a representaciones proposicionales; se procesarían también en el “mentales”. Pero existen otros, “los imagistas”, que no aceptan esta posición y defienden una teoría de imágenes según la cual dichas representaciones tienen propiedades especiales y no pueden ser reducidas a proposiciones. Argumentan, por ejemplo, que las imágenes pueden ser rotadas, trasladadas y escudriñadas mentalmente. Por lo tanto, tienen identidad propia, son una forma representacional independiente.

La cuestión imágenes/proposiciones es polémica en la Psicología Cognitiva. Hay defensores férreos de ambas posiciones. Pero hay también una tercera vía, una síntesis, una tercera forma de constructo representacional, llamada *modelos mentales*, propuesta por Johnson-Laird (1983).

Para él, las *proposiciones* son representaciones de significados, totalmente abstraídas, que *son verbalmente expresables*. El criterio de expresabilidad verbal distingue a Johnson-Laird de otros psicólogos cognitivos (Sternberg, 1996, p. 181). Las *imágenes* son representaciones bastante específicas que retienen muchos de los aspectos perceptivos de determinados objetos o eventos, vistos desde un ángulo particular, con detalles de una cierta instancia del objeto o evento. Los *modelos mentales* son representaciones analógicas, algún tanto abstraídas, de conceptos, objetos o eventos que son espacial y temporalmente análogos a impresiones sensoriales, pero que pueden ser vistas desde cualquier ángulo (¡y ahí tenemos imágenes !) y que, en general, no retienen aspectos distintivos de una instancia dada de un objeto o de un evento (ibid.).

Por ejemplo, la situación “el cuadro está en la pared” podría representarse mentalmente como una proposición (porque es verbalmente expresable), como un modelo mental (de cualquier cuadro en cualquier pared, posiblemente prototípicos) o como una imagen (de un cuadro en particular en cierta pared) .

### **Modelos mentales, proposiciones e imágenes**

Johnson-Laird sugiere que las personas razonan con modelos mentales. Los modelos mentales son como bloques de construcción cognitivos que pueden ser combinados y recombinados conforme sea necesario. Como cualesquiera otros modelos, éstos representan el objeto o situación en sí ; una de sus características más importantes es que su estructura capta la esencia (se parece analógicamente) de esa situación u objeto (Hampson y Morris, 1996, p. 243).

*Un modelo mental es una representación interna de informaciones que corresponde análogamente a aquello que se está representando.*

La analogía puede ser total o parcial, o sea, un modelo mental es una representación que puede ser totalmente analógica o parcialmente analógica y parcialmente proposicional (Eisenck y Keane, 1994, p. 209). Es decir, un modelo mental puede contener proposiciones, pero éstas pueden existir como representación mental, en el sentido de Johnson-Laird, sin formar parte de un modelo mental. No obstante, para él, las representaciones proposicionales se interpretan en relación con modelos mentales : una proposición es verdadera o falsa en relación con un modelo mental de un estado de cosas del mundo. Las imágenes, a su vez, corresponden a visiones de los modelos.

Por lo tanto, en la perspectiva de Johnson-Laird, *las representaciones proposicionales son cadenas de símbolos que corresponden al lenguaje natural, los modelos mentales son análogos estructurales del mundo y las imágenes son modelos vistos desde un determinado punto de vista* (1983, p. 165).

Las *imágenes*, para Johnson-Laird son producto tanto de la percepción como de la imaginación. Representan aspectos perceptibles de los objetos correspondientes en el mundo real. En un primer momento correspondían para el autor a vistas concretas de un modelo mental subyacente, a visiones relevantes o “visuales” del mismo; pero en un trabajo más reciente el propio Johnson-Laird (1996) revisa esta idea diferenciando las imágenes de los modelos mentales y no considerándolas solamente como vistas de un modelo mental que actúe como sustrato. “*Las imágenes representan cómo algunas cosas son vistas desde un punto de vista particular*” (Johnson-Laird, 1996, p. 124) y no suponen necesariamente la construcción o la posesión de un modelo mental explicativo y predictivo del que deriven. Esta revisión parece pertinente y quizás un ejemplo claro pueda ser el concepto “célula” para el que muchos estudiantes generan una imagen simple, estática, “huevo frito”, que opera en sus mentes de manera aislada, proposicionalmente, pero a la que no le atribuyen ningún sentido, ningún significado, no suponiendo su posesión la construcción

de un modelo mental subyacente como análogo de lo que una célula es conceptualmente (Rodríguez Palmero, 1999).

Según él, los modelos mentales y las imágenes son representaciones de alto nivel, esenciales para el entendimiento de la cognición humana (Eisenck y Keane, 1994, p. 210). Aunque en su nivel básico el cerebro humano pueda computar las imágenes y los modelos mentales en algún código proposicional (el “mentalés”), el uso de estas representaciones libera a la cognición de la obligación de operar proposicionalmente en “código de máquina”. Estas representaciones de alto nivel pueden compararse con las de los lenguajes de programación de los ordenadores. En última instancia, el ordenador trabaja en un código binario, pero el programador no : él usa lenguajes de alto nivel que le permiten pensar sobre lo que el ordenador tiene que hacer usando el código binario. Los lenguajes de programación de alto nivel son traducidos por los ordenadores a códigos binarios cuando se compilan. Análogamente, las imágenes y los modelos mentales podrían traducirse en la mente en algún código proposicional semejante al código binario. La metáfora del ordenador, la mente como un sistema de cómputo, es un credo fundamental de la psicología cognitiva, pero eso no significa que la mente opere necesariamente en un código binario. La mente tiene un código propio, el “mentalés”, que no es consciente, al cual no tenemos acceso y no precisamos tenerlo pues operamos muy bien con proposiciones, imágenes y modelos mentales (todos en el sentido de Johnson-Laird).

## **Modelos mentales**

Supongamos que a un grupo de personas se le dé un conjunto bien determinado de descripciones de una distribución espacial (indicando la posición exacta de cada objeto en la composición espacial) y a otro grupo de personas se le dé un conjunto que no esté bien determinado de descripciones de la misma organización espacial (dando localizaciones ambiguas, poco precisas, de los objetos en el arreglo espacial).

Mani y Johnson-Laird (1982, apud. Sternberg, 1996, p. 181) hicieron una investigación de este tipo y encontraron que los sujetos que recibieron informaciones bien determinadas fueron capaces de inferir informaciones espaciales adicionales no incluidas en las descripciones que recibieron, pero tuvieron dificultades para recordar literalmente las informaciones recibidas. Esta constatación se interpretó como indicadora de que esos sujetos formaron un modelo mental de la información recibida y, por eso mismo, fueron capaces de hacer inferencias. Dicho esto, por haber formado el modelo pasaron a confiar en él en vez de quedarse dependiendo de recordar descripciones verbales literales detalladas.

Por otro lado, los sujetos que recibieron informaciones poco precisas raramente fueron capaces de inferir informaciones espaciales no incluidas en las descripciones recibidas ; sin embargo, recordaban mejor que los del otro grupo estas descripciones. Los investigadores sugirieron que en este caso los sujetos no construyeron un modelo mental debido a las innumerables posibilidades de modelos mentales que podrían inferirse a partir de las informaciones (indeterminadas) recibidas. En vez de eso, los sujetos parecen haber representado mentalmente las descripciones recibidas como proposiciones verbalmente expresables (ibid.).

Es decir, en ambos casos los sujetos representaron mentalmente el arreglo espacial, pero en el primero formaron un modelo mental que les permitió hacer inferencias y en el segundo trabajaron sólo con un conjunto de proposiciones descriptivas.

*Los modelos mentales son, por lo tanto, una forma de representación analógica del conocimiento : existe una correspondencia directa entre entidades y relaciones presentes en la estructura de esa representación y las entidades y relaciones que se quieren representar.*

Un modelo mental está compuesto por *elementos* (“tokens”)<sup>2</sup> y *relaciones* que representan un estado de cosas específico, estructurados de una manera adecuada al proceso sobre el que deberán operar. O sea, cada modelo ya se construye de una manera coherente con el uso previsto (STAF 11, 1996).

No existe un único modelo mental para un determinado estado de cosas. Al contrario, pueden existir varios, incluso aunque sólo uno de ellos represente de manera óptima ese estado de cosas (ibid.). Cada modelo mental es *una* representación analógica de ese estado de cosas y, recíprocamente, cada representación analógica corresponde a *un* modelo mental (ver p. 11, *principio de la economía*).

Los estados de cosas muchas veces se describen por conceptos. El modelo mental de un concepto debe ser capaz de representar tanto lo esencial como la amplitud de ese concepto. El núcleo del modelo representa lo esencial del concepto, o sea, las propiedades características del estado de cosas que describe ; los procedimientos de gestión del modelo definen la amplitud de ese concepto, es decir, el conjunto de estados de cosas posibles que el concepto describe (ibid.).

El modelo mental de avión, por ejemplo, posee distintas versiones conforme los diferentes usos que se puedan hacer de un avión : reconocerlo, construirlo, pilotarlo, embarcar en él, hablar sobre él. El modelo varía también según otras dimensiones : la competencia aeronáutica del sujeto, su edad, su cultura, etc. Representar un avión en vuelo o un avión abierto para mostrar los lugares a los pasajeros también corresponde a diferentes versiones del modelo mental de avión. Cada versión, sin embargo, debe incluir el núcleo central que identifica el modelo como modelo de un avión. Debe también incluir proposiciones y procedimientos de manipulación diversificados, teniendo en cuenta que, en función del uso, son otros los aspectos del modelo que son accionados. Es posible que dos ejemplares del mismo modelo poco o nada tengan en común si se construyeron con finalidades totalmente diferentes (ibid.).

En este punto es conveniente anticipar la distinción hecha por Norman (en Gentner y Stevens, 1983, p. 8) entre modelos conceptuales y modelos mentales, que se discutirá más adelante : *los modelos conceptuales se proyectan como instrumentos para la comprensión o para la enseñanza de sistemas físicos ; los modelos mentales son lo que las personas realmente tienen en sus cabezas y lo que guía el uso que hacen de las cosas*. Idealmente, debería haber una relación directa y simple entre el modelo conceptual y el modelo mental. Muy frecuentemente, sin embargo, no es esto lo que sucede.

Los modelos a los que se refiere Johnson-Laird, de los que hemos estado hablando hasta aquí (incluso en el caso del modelo del concepto de avión) son, por lo tanto, mentales y pueden no tener una relación directa y simple con algún modelo conceptual en el significado dado por Norman. ¡Es importante que esto quede claro !.

Repetimos : *modelo mental es una representación interna de informaciones que corresponde, análogamente, al estado de cosas que se están representando, sea cual sea el mismo. Los modelos mentales son análogos estructurales del mundo.*

---

<sup>2</sup> “Token”, en el sentido empleado en la teoría de los modelos mentales podría significar señal, signo, indicio, rasgo característico. Sin embargo, creemos que ninguno de esos significados es adecuado como traducción de “token” al castellano en el contexto de esa teoría. Optamos, entonces, por el término “elemento”, ante la falta de otro mejor.

## Modelos mentales y razonamiento deductivo

Para Johnson-Laird, en vez de una lógica mental, las personas usan modelos mentales para razonar. Los modelos mentales, como ya se ha dicho, son como bloques de construcción cognitivos que pueden combinarse y recombinarse conforme sea necesario. El aspecto esencial del razonamiento a través de modelos no está sólo en la construcción de modelos adecuados para captar distintos estados de cosas, sino también en la habilidad de verificar cualesquiera conclusiones a las que se llegue usando tales modelos. La lógica, si es que aparece en algún lugar, no está en la construcción de modelos sino en la verificación de las conclusiones pues ésta implica que el sujeto sepa apreciar la importancia lógica de falsear una conclusión y no sólo buscar evidencia positiva que la apoye (Hampson y Morris, 1996, p. 243).

En esta perspectiva, el razonamiento deductivo es mejor interpretado como una destreza práctica que como una habilidad esotérica, abstracta. Dicho esto, lo que separaría a “especialistas y novatos”, “experimentados y principiantes”, en términos de razonamiento, serían diferencias en el espacio disponible en la memoria de trabajo para construir y manipular modelos mentales complejos, así como la persistencia en la verificación de conclusiones (aunque el éxito en tareas de razonamiento dependa también del conocimiento y la experiencia del individuo) (ibid.).

En la teoría de Johnson-Laird están ausentes las reglas de inferencia de la lógica formal. Por eso, es considerada no-racionalista. En ella, la resolución de tareas de inferencia silogística está basada en la manipulación de modelos mentales, no en la lógica formal. Un ejemplo adaptado de otro dado por el propio Johnson-Laird (1981 ; apud de Vega, 1984, p. 453-454) es el que considera las siguientes premisas :

- Todos los investigadores son profesores.
- Todos los licenciados son profesores.

La elaboración de modelos mentales que satisfagan estas premisas puede llevarse a cabo en un escenario hipotético, en el que se disponga de actores que desempeñen papeles de investigador, licenciado y profesor, estableciéndose relaciones adecuadas. Tal elaboración puede llevar a muchos sujetos a conclusiones equivocadas como la de que “todos los investigadores son licenciados”. Pero un razonador cuidadoso seguirá elaborando modelos alternativos, con la intención de falsear o confirmar la conclusión, observando siempre las premisas. En este proceso puede concluirse que “algunos investigadores son licenciados”, pero tal conclusión debe verificarse y puede ser falseada por un nuevo modelo en el cual todos los investigadores son profesores, pero ninguno es licenciado. La consecuencia, ahora correcta, es que de esas premisas no se puede sacar ninguna conclusión válida que relacione los papeles de investigador y licenciado.

Consideremos un ejemplo más, adaptado de otro dado por Hampson y Morris (1996, p. 243-244). Supongamos los siguientes enunciados :

- El lápiz está a la izquierda del portaplumas.
- La goma está enfrente del portaplumas.
- La regla está enfrente del lápiz.

Sin el uso de la lógica formal, se puede construir un modelo mental que capta la composición espacial de esos objetos :

lápiz  
regla

portaplumas  
goma

Examinando este modelo se puede sacar la conclusión simple y no ambigua de que “la regla está a la izquierda de la goma”.

Obviamente, cuanto más complicadas sean las proposiciones originales, más difícil será la construcción y el mantenimiento de un modelo integrado. Dicho esto, en algunos casos la combinación de enunciados puede admitir más de una interpretación.

Supongamos que los enunciados son los siguientes :

El lápiz está a la izquierda del portaplumas.  
La goma está a la izquierda del portaplumas.

Rápidamente se puede formar un modelo mental:

lápiz

goma

portaplumas

y concluir que la goma está a la derecha del lápiz.

Sin embargo, una persona más experimentada en razonamiento intentará falsear esta conclusión buscando otro modelo que satisfaga los enunciados. Ese modelo podría ser :

goma

lápiz

portaplumas

La existencia de dos modelos posibles significa, en este caso, que no existe una única conclusión no ambigua que se pueda sacar de la relación entre lápiz y goma a partir de las proposiciones iniciales.

Está claro que para Johnson-Laird la teoría de los modelos mentales se aplica a mucho más que a las distribuciones espaciales irrefutables como las de este ejemplo. Para él, la teoría da cuenta también de razonamientos silogísticos abstractos que incluyen términos como “todos”, “ningún” y “algunos”.

Según de Vega (1984, p. 454), el procedimiento de elaboración de inferencias silogísticas de Johnson-Laird implica tres etapas :

- 1.- construir un modelo mental de la primera premisa;
- 2.- agregar la información de la segunda premisa al modelo mental de la primera, teniendo en cuenta los modos alternativos en los que esto puede hacerse;
- 3.- inferir una conclusión que exprese la relación, si existiera, entre los términos extremos, que sea común a todos los modelos de las premisas construidos en las etapas previas.

En el ejemplo de los investigadores y licenciados no existe esta relación entre los términos extremos, investigador y licenciado, común a todos los modelos y, por lo tanto, no se deriva ninguna conclusión de interés.

Etapa 1 :

investigador = profesor  
 profesor (profesor no investigador)

Etapa 2 :

investigador = profesor = licenciado  
 profesor (profesor no investigador y no licenciado).

Modelo alternativo :

investigador = profesor = licenciado  
 investigador = profesor (investigador no licenciado)  
 profesor = licenciado (licenciado no investigador)

Modelo alternativo :

investigador = profesor  
 profesor = licenciado

Según Johnson-Laird, las dificultades de muchos problemas de razonamiento deductivo están relacionadas con el número de modelos mentales necesarios para representar adecuadamente las premisas del argumento deductivo. Argumentos que implican sólo un modelo mental pueden resolverse rápida y cuidadosamente. Sin embargo, es muy difícil sacar conclusiones precisas basadas en argumentos que pueden ser representados por múltiples modelos alternativos debido a la gran demanda hecha sobre la memoria de trabajo. El individuo en este caso debe mantener en la memoria de trabajo cada uno de los varios modelos para poder llegar a una conclusión o para verificar una conclusión (Sternberg, 1996, p. 410). Una manera de perfilar esa limitación de la memoria de trabajo es representar la información implícitamente el máximo posible, en vez de explícitamente.

## Modelos conceptuales y modelos mentales

Los *modelos conceptuales* se han inventando por los profesores, investigadores, ingenieros, arquitectos, para facilitar la comprensión o la enseñanza de sistemas físicos o estados de cosas físicos. Son representaciones precisas, consistentes y completas de sistemas físicos. Se proyectan como herramientas para la comprensión o para la enseñanza de sistemas físicos (Norman, apud. Gentner y Stevens, 1983, p. 7).

Los *modelos mentales* son modelos que las *personas* construyen para representar estados físicos. Esos modelos no precisan ser técnicamente precisos (y generalmente no lo son), sino que deben ser funcionales. Evolucionan naturalmente. Interactuando con el sistema, la persona continuamente modifica su modelo mental con el fin de llegar a una funcionalidad que le satisfaga. Está claro que los modelos mentales de una persona están limitados por factores tales como su conocimiento y su experiencia previa con sistemas similares y por la propia estructura del sistema de procesamiento humano de la información (op. cit. p. 8).

Norman (ibid.) sugiere que los modelos mentales tienen las siguientes características generales :

- 1.- los modelos mentales son incompletos ;
- 2.- la habilidad de las personas para ejecutar (“rodar”) sus modelos mentales es muy limitada ;
- 3.- los modelos mentales son inestables : las personas olvidan detalles del sistema modelado, particularmente cuando esos detalles (o todo el sistema) no son utilizados por un cierto periodo de tiempo ;

4.- los modelos mentales no tienen fronteras bien definidas : dispositivos y operaciones similares se confunden unos con otros ;

5.- los modelos mentales son “no-científicos” : las personas mantienen patrones de comportamiento “supersticiosos”, incluso cuando saben que no son necesarios (por ejemplo, apretar la tecla CLEAR o la tecla ENTER, de una calculadora varias veces, innecesariamente, “solo para tener la certeza”) ; los modelos mentales de una persona reflejan sus creencias sobre el sistema físico ;

6.- los modelos mentales son parsimoniosos ; frecuentemente las personas optan por operaciones físicas adicionales en vez de un planteamiento mental que evitaría tales operaciones ; las personas prefieren gastar más energía física a cambio de menor complejidad mental.

La principal función del modelo mental es permitir a su constructor explicar y hacer previsiones sobre el sistema físico que el modelo analógicamente representa. Tales previsiones no implican necesariamente “rodar” el modelo (previsibilidad procedimental), pues debe también permitir previsiones resultantes de inferencia directa (previsibilidad declarativa) (op. cit., p. 13).

¿Pero cuál es la relación entre modelos conceptuales y modelos mentales ? Idealmente debería haber una relación directa y simple entre ambos. En la práctica, no obstante, no es éste el caso.

Los modelos conceptuales son delineados, proyectados, por las personas que usan modelos mentales, para facilitar la comprensión de sistemas físicos por parte de otras personas que también utilizan modelos mentales. En la enseñanza, el profesor enseña modelos conceptuales y espera que el estudiante construya modelos mentales, robustos y acordes con esos modelos conceptuales que, a su vez, deben ser consistentes con los sistemas físicos modelados. Los modelos conceptuales son, por tanto, instrumentales, medios y no fines. El objetivo de la enseñanza es, a través de modelos conceptuales, llevar al aprendiz a formar modelos mentales adecuados (i.e., consistentes con los propios modelos conceptuales) de los sistemas físicos. Es decir, la mente humana opera sólo con modelos mentales , pero los modelos conceptuales pueden ayudar a la construcción de modelos mentales que explican y predicen de acuerdo con el conocimiento aceptado en una área dada.

Para eso, los modelos conceptuales enseñados deben, según Norman (ibid.) ser *aprendibles, funcionales y utilizables*.

Estos criterios parecen obvios (¿para qué serviría un modelo conceptual que es difícil en demasía para aprenderlo ? ¿o un modelo que no explica o predice importantes aspectos del sistema físico ? ¿o que no puede utilizarse fácilmente dentro de la limitada memoria de trabajo o limitada capacidad computacional humanas ?), pero no siempre se observan en la enseñanza.

¿Y la cuestión de modelar los modelos mentales de una persona ? ¿Cómo identificar el modelo mental que una persona formó para un determinado estado de cosas, cierto sistema físico, por ejemplo ? Ciertamente es necesario tener un modelo conceptual de ese estado de cosas, pero es preciso también distinguir entre el modelo de modelo mental (la conceptualización de modelo mental) de quien investiga y el modelo mental que se piensa que la persona, cuyos modelos se investiga, de hecho tiene.

Norman (op. cit., p. 12) sugiere tres factores funcionales que se aplican tanto al modelo mental como al modelo conceptual de modelo mental :

*El sistema de creencias.* Los modelos mentales de una persona reflejan sus creencias sobre el sistema físico representado, adquiridas por observación, instrucción o inferencia. El modelo conceptual de modelo mental debe contemplar el sistema de creencias de la persona.

*Observabilidad.* Debe haber una correspondencia entre parámetros y estados del modelo mental que son accesibles a la persona y parámetros y estados del sistema físico que la persona puede observar. Esta misma correspondencia debe existir entre parámetros y estados del modelo conceptual de modelo mental y el sistema físico.

*Potencia predictiva.* La finalidad de un modelo mental es permitir que la persona entienda y anticipe el comportamiento del sistema físico. Eso significa que el modelo debe tener una potencia predictiva tanto a través de reglas de inferencia como por medio de derivaciones procedimentales ; o sea, la persona debe poder “rodar” mentalmente el modelo. Por tanto, el modelo conceptual de modelo mental debe considerar también las estructuras de conocimiento y de procesamiento de información humanas que le permiten a la persona usar su modelo mental para comprender y prever el sistema físico.

Resumiendo esta sección :

1. Es preciso distinguir entre sistema físico, modelo conceptual del sistema físico y modelo mental del sistema físico. El modelo conceptual es un modelo preciso, consistente y completo del sistema físico que se inventa para facilitar la construcción de un modelo mental (que no es preciso, consistente y completo, pero debe ser funcional) adecuado (con poder explicativo y predictivo) del sistema físico. Es importante advertir que los modelos conceptuales son inventados por las personas que operan mentalmente con modelos mentales. Es también importante observar que para identificar modelos mentales de otras personas es preciso tener un modelo de modelo mental, es decir, un modelo conceptual de modelo mental.

Los modelos mentales de las personas pueden ser deficientes en varios aspectos, tal vez, incluyendo elementos innecesarios, erróneos o contradictorios. En la enseñanza, es preciso desarrollar modelos conceptuales y también materiales y estrategias instruccionales que ayuden a los estudiantes a construir modelos mentales adecuados. En la investigación, es necesario desarrollar técnicas de investigación apropiadas y, en vez de buscar modelos mentales claros y elegantes, procurar entender los modelos confusos, “desordenados”, incompletos, inestables, que las personas realmente tienen.

## **Naturaleza de los modelos mentales**

De todo lo que se ha dicho hasta aquí, debe haber quedado claro que es difícil definir e identificar exactamente lo que son los modelos mentales y cómo difieren de otras formas postuladas de representaciones mentales como los esquemas de Piaget, los subsumidores de Ausubel y los constructos personales de Kelly. Johnson-Laird reconoce eso (1983, cap. 15), pero en vez de seguir diferenciando directamente el concepto de modelo mental, distinguiéndolo explícitamente de propuestas de otros autores, prefiere apuntar una serie de principios que imponen vínculos a la naturaleza de los modelos mentales y limitan tales modelos (op. cit., p. 398).

1.- *Principio de la computabilidad* : los modelos mentales son computables, i.e., deben poder describirse en la forma de procedimientos efectivos que puedan ser ejecutados por una máquina. (Este vínculo viene del “núcleo duro” de la Psicología Cognitiva que supone la mente como un sistema de cómputo). Procedimiento efectivo es aquel que puede llevarse a cabo sin implicar ninguna decisión basada en la intuición o cualquier otro ingrediente “misterioso” o “mágico”.

2.- *Principio de lo finito* : los modelos mentales son finitos en tamaño y no pueden representar directamente un dominio infinito. Este vínculo deriva de la premisa de que el cerebro es un órgano finito.

3.- *Principio del constructivismo* : los modelos mentales se construyen a partir de elementos básicos (“tokens”) organizados en una cierta estructura para representar un determinado estado de cosas. Este vínculo surge de la función primaria de los modelos mentales que es la de representar mentalmente estados de cosas. Como existe un número infinito de estados de cosas que puede representarse pero solamente un mecanismo finito para construir modelos que los representen, se deriva que tales modelos deben construirse a partir de constituyentes más elementales.

4.- *Principio de la economía* : una descripción de un único estado de cosas es representada por un único modelo mental, incluso si la descripción es incompleta o indeterminada. Pero un único modelo mental puede representar un número infinito de posibles estados de cosas porque ese modelo puede revisarse recursivamente. Cada nueva aseveración descriptiva de un estado de cosas puede implicar revisión del modelo para acomodarla. Este vínculo se refiere particularmente a la construcción de modelos a partir del discurso, pues éste es siempre indeterminado y compatible con muchos estados de cosas diferentes ; para perfilar esto, la mente construye un modelo mental inicial y lo revisa recursivamente conforme sea necesario. Naturalmente hay límites para esa revisión : en última instancia, el proceso de revisión recursiva es gobernado por las condiciones de verdad del discurso en el que el modelo se basa (op. cit., p. 408).

5.- *Principio de la no-indeterminación* : los modelos mentales pueden representar indeterminaciones directamente si y sólo si su uso no fuera computacionalmente intratable, i.e., si no existiera un crecimiento exponencial en complejidad. Este vínculo es un corolario del primero y del anterior : si se tratara de acomodar cada vez más indeterminaciones en un modelo mental, eso llevará rápidamente a un crecimiento intratable en el número de posibles interpretaciones del modelo que, en la práctica, dejará de ser un modelo mental (op. cit. p. 409).

6.- *Principio de la predicabilidad* : un predicado puede ser aplicable a todos los términos a los que otro predicado es aplicable, pero no pueden tener ámbitos de aplicación que no se intersecten. Por ejemplo, los predicados “animado” y “humano” son aplicables a ciertas cosas en común, “animado” se aplica a algunas cosas a las que “humano” no se aplica, pero no existe nada a lo que se aplique “humano” y “animado” no. Para Johnson-Laird (p. 411), la virtud de este vínculo es que permite identificar un concepto artificial o no natural. Un concepto que se definiese por predicados que no tuvieran nada en común violaría el principio de predicabilidad y no estaría, normalmente, representado en modelos mentales.

7.- *Principio del innatismo* : todos los primitivos conceptuales son innatos. Los primitivos conceptuales subyacen a nuestras experiencias perceptivas, habilidades motoras, estrategias, en fin, nuestra capacidad de representar el mundo (ibid.). Indefinibilidad es una condición suficiente, pero no necesaria, para identificar conceptos primitivos. Movimiento, por ejemplo, es una palabra que corresponde a un primitivo conceptual, pero que puede definirse. Aunque proponga este vínculo a los modelos mentales, Johnson-Laird rechaza el innatismo extremo de que todos los conceptos son innatos aunque algunos tengan que ser “disparados” por la experiencia. Él defiende el aprendizaje de conceptos a partir de primitivos conceptuales innatos o de conceptos previamente adquiridos (p. 412). Además de los primitivos conceptuales innatos, admite también la existencia de primitivos procedimentales que se accionan automáticamente cuando un individuo construye un modelo mental. Los primitivos procedimentales no pueden adquirirse a través de la experiencia porque la representación mental de la experiencia ya requiere la habilidad de construir modelos de la realidad a partir de la percepción. Estos primitivos deben ser innatos (op. cit. p. 413).

8.- *Principio del número finito de primitivos conceptuales* : existe un conjunto finito de primitivos conceptuales que origina un conjunto correspondiente de campos semánticos y otro conjunto finito de conceptos, u “operadores semánticos”, que se da en cada campo semántico y sirve para construir conceptos más complejos a partir de los primitivos subyacentes. Un campo semántico se refleja en el léxico por un gran número de palabras que comparten en el núcleo de sus significados un concepto común. Por ejemplo, los verbos asociados a la percepción visual como avistar, ojear, escrutar y observar comparten un núcleo subyacente que corresponde al concepto de ver. Los operadores semánticos incluyen los conceptos de tiempo, espacio, posibilidad, permisibilidad, causa e intención. Por ejemplo, si las personas ojean alguna cosa, enfocan sus ojos durante cierto intervalo de tiempo con la intención de ver lo que ocurre. Los campos semánticos proveen nuestra concepción sobre lo que existe en el mundo, sobre el mobiliario del mundo, mientras que los operadores semánticos proveen nuestro concepto sobre las posibles relaciones que pueden ser inherentes a esos objetos (p. 414).

9.- *Principio de la identidad estructural* : las estructuras de los modelos mentales son idénticas a las estructuras de los estados de cosas, percibidos o concebidos, que los modelos representan. Este vínculo deriva, en parte, de la idea de que las representaciones mentales deben ser económicas y, por lo tanto, cada elemento de un modelo mental, incluyendo sus relaciones estructurales, debe tener un papel simbólico. No debe haber en la estructura del modelo ningún aspecto sin función o significado (p. 419).

### **Estructura y contenido de los modelos mentales**

A diferencia de las representaciones proposicionales, los modelos mentales no tienen una estructura sintáctica ; su estructura es análoga a la estructura de los estados de cosas del mundo que esos modelos representan, tal como los percibimos o concebimos; los modelos mentales son análogos estructurales de estados de cosas del mundo (p. 156). No obstante, la estructura analógica de los modelos mentales puede variar bastante : modelos contruidos a partir de proposiciones discretas pueden tener sólo un mínimo de estructura analógica, mientras que modelos mentales de arreglos espaciales, como un laberinto, por ejemplo, pueden tener una gran analogía estructural en dos, tres o quién sabe cuantas dimensiones (ibid.).

Las representaciones proposicionales pueden ser escudriñadas solamente en las direcciones permitidas por la sintaxis y por las codificaciones de ese tipo de representación. Los modelos mentales, por su estructura dimensional, pueden manipularse más libremente, de manera controlada sólo por las propias dimensiones del modelo. Las imágenes, como ya se ha destacado anteriormente, corresponden a *visiones* de modelos : resultantes de percepción o de imaginación, éstas representan aspectos perceptibles de los objetos o eventos correspondientes del mundo real (p. 157).

En términos de *contenido*, los modelos mentales, las imágenes y las proposiciones presentan una diferencia importante en lo que se refiere a la especificidad : los modelos mentales, así como las imágenes, son altamente específicos. Por ejemplo, no es posible formar una imagen de un objeto (un cuadro, una mesa o avión) *en general* sino de un *objeto específico* (un determinado cuadro, mesa o avión). Las representaciones proposicionales, sin embargo, no implican tanta especificidad : es perfectamente aceptable, por ejemplo, una representación mental proposicional que establezca la relación espacial entre dos objetos como si estuvieran “al lado de”, sin explicitar “izquierda” o “derecha”. Para una imagen eso no sería posible (p. 158).

Aun en lo que se refiere al contenido de los modelos mentales, Johnson-Laird (op. cit. p. 410) dice que “como pueden tener muchas formas y servir para muchas finalidades, sus contenidos

son muy variados : pueden contener solamente elementos que representan individuos e identidades entre ellos, como en los modelos necesarios en el razonamiento silogístico ; pueden representar relaciones espaciales entre entidades o relaciones temporales o causales entre eventos. Los modelos mentales tienen el contenido y la forma que sirven a las finalidades para las cuales fueron contruidos, sean explicar, predecir o controlar”. La estructura de los modelos corresponde a la estructura de los estados de cosas del mundo, tal como son percibidos o concebidos por el individuo, así representados.

*La naturaleza de los modelos mentales es, por lo tanto, más restringida que sus contenidos.* En la medida en que los modelos mentales son análogos estructurales del mundo, su estructura está limitada por la estructura de los estados de cosas del mundo. La cuestión de los contenidos, por otro lado, es ontológica (del ser humano como ser humano) : nuestra concepción de lo que existe es función del mundo y de nuestra capacidad de concebir que sería, en principio, ilimitada. Pero Johnson-Laird argumenta (ibid) que hay límites y que estarían en los *conceptos* que subyacen a los significados de las cosas, una vez que los conceptos son restringidos por la naturaleza del aparato cognitivo humano. Los principios de la predicabilidad, del innatismo y del número finito de primitivos conceptuales, introducidos en la sección anterior, son los tres principales vínculos que él dice que afectan a los contenidos posibles de los modelos mentales. El primero de ellos explica por qué ciertos conceptos no son naturales y normalmente no aparecen en los modelos mentales ; el segundo y el tercero limitan el contenido de los modelos mentales en términos de sus componentes básicos y de cómo pueden organizarse. O sea, existe un conjunto finito de primitivos conceptuales, un conjunto finito correspondiente de campos semánticos y otro conjunto finito de operadores semánticos que imponen límites a los modelos posibles.

### **Modelos mentales según otros autores**

En este punto, es interesante ver lo que otros autores dicen sobre la naturaleza y contenido de los modelos mentales. Para Williams, Hollan y Stevens (1983, p. 133) los modelos mentales están compuestos de *objetos autónomos* con una cierta topología, son “*rodables*” por medio de inferencias cualitativas locales y *pueden descomponerse*. La noción de objeto autónomo es central en su concepción de modelos mentales : se trata de un objeto mental que representa explícitamente alguna cosa, cuyas conexiones topológicas con otros objetos son también explícitas, y que tiene un cierto número de parámetros internos. Asociado a cada objeto autónomo, existe un conjunto de reglas (internas) para modificar sus parámetros y, así, especificar su contenido.

Para estos autores, un modelo mental es, entonces, un conjunto de objetos autónomos “conectados”. Por ejemplo, (op. cit. p. 134), una “región de fluido” puede funcionar como un objeto autónomo en un modelo y tener como parámetro la temperatura, que puede asumir uno de cuatro “valores” (aumentando, disminuyendo, constante o indeterminada). Este objeto está conectado, explícitamente, con un número limitado de otros objetos e interactúa con ellos transmitiendo cambios en los valores de sus parámetros a través de lo que los autores llaman “puertos” (conexiones mentales entre los objetos autónomos que componen el modelo mental). Así, el modelo mental de un “cambiador de calor” (“heat exchanger”) puede incluir un objeto autónomo que representa el mecanismo de transferencia de energía térmica y varios objetos autónomos que representan “regiones de fluido”. Algunos de estos objetos autónomos estarán conectados a lo que corresponde al mecanismo de transferencia y solamente ellos podrán pasar hacia adelante la información de que hubo un cambio en los parámetros de cierta región de fluido. Es esa propagación de los cambios en los valores de los parámetros lo que da el sentido de “rodar” el modelo, siempre presente en el razonamiento sobre sistemas físicos.

Los objetos autónomos de Williams, Hollan y Stevens son objetos mentales (que posiblemente serían también modelos mentales) que tienen fronteras definidas. Su comportamiento, *definido como cambios en los valores de los parámetros*, es gobernado por operaciones (reglas) internas no directamente accesibles, no significativas. El resultado de eso es una cierta opacidad del objeto autónomo de modo que no se podría preguntar cómo se produce un determinado comportamiento, pero sí observarlo.

Sin embargo, aunque los objetos autónomos sean normalmente opacos, pueden a veces, descomponerse. El resultado de esa descomposición es un nuevo modelo mental, compuesto también por objetos autónomos con una topología dada, que pueden usarse para producir explicaciones del comportamiento del objeto autónomo de nivel más alto (op. cit. p. 135). Este nuevo modelo estaría impregnado, encajado, incrustado, en el modelo original.

Eso significaría que el “funcionamiento” de un modelo mental podría explicarse por otro modelo mental, resultante de la descomposición del anterior, que sería subyacente. Este modelo subyacente podría usarse para inferir el comportamiento (cambios en los valores de los parámetros) de un objeto autónomo en condiciones no especificadas (tal vez, olvidadas) en el funcionamiento del modelo inicial de “nivel más alto”. En esta línea de razonamiento, el nuevo modelo podría, en principio, descomponerse generando otro modelo subyacente de “nivel más bajo”. El límite inferior de este proceso sería, tal vez, función de aquellos conjuntos finitos de primitivos conceptuales, de campos semánticos y de operadores semánticos de los que habla Johnson-Laird.

Otros autores cuyo trabajo en el área de los modelos mentales ha sido muy citado son de Kleer y Brown (1983). Su objetivo es desarrollar modelos sobre la manera en la que el individuo comprende el funcionamiento de dispositivos tales como máquinas, aparatos electrónicos, hidráulicos, térmicos, ... . A estos modelos les dan el nombre de *modelos mentales mecanicistas* (op. cit. p. 155). Su construcción implica cuatro aspectos relacionados : el más básico es la *topología del dispositivo* (una representación de su estructura, su organización física, sus componentes) ; el segundo es un proceso de inferencia, llamado “*visionamiento*” (“envisioning”), que a partir de la estructura (topología) del dispositivo determina su función ; el tercero, denominado *modelo causal*, describe el funcionamiento del dispositivo (i.e., una descripción de la forma por la que el comportamiento del dispositivo deriva de interacciones causales entre los componentes) ; el último es la *ejecución* (“rodaje”) del modelo causal, a través de una serie de eventos cada uno de ellos relacionado causalmente con el anterior, para producir un comportamiento específico del dispositivo (p. 158).

El modelo de de Kleer y Brown es aparentemente simple, pero, basados en sus investigaciones, ellos dicen que es sorprendentemente difícil construir modelos mentales de dispositivos electro-mecánicos, por ejemplo, si estos modelos deben ser capaces de prever consecuencias de eventos que no se consideraron en la creación del modelo. Para ellos, el proceso de construcción de un modelo mecanicista implica dos problemas distintos : uno es la construcción de una simulación cualitativa mental del dispositivo y el otro es la simulación mental del resultado del funcionamiento de esta construcción ; el primero de estos problemas está relacionado con el proceso que ellos llaman “visionamiento” y el segundo con lo que denominan ejecución (“rodaje”) del modelo causal.

Idealmente, un modelo causal debe ser *consistente, correspondiente y robusto* (p. 167). Un modelo causal consistente es aquel que no tiene contradicciones internas : distintos componentes no pueden dar valores diferentes para un mismo atributo de un estado del dispositivo. *Correspondencia* significa que el modelo causal debe ser fiel al comportamiento real del dispositivo modelado. *Robustez* tiene que ver con la utilidad del modelo causal en situaciones atípicas, por ejemplo, cuando la estructura del dispositivo ha sido perturbada. Un modelo robusto debe prever el

comportamiento del dispositivo cuando, por ejemplo, una de sus partes es defectuosa. Según estos autores, la mejor manera de tener un modelo causal robusto es tener robustos los propios componentes del modelo.

Más adelante se darán ejemplos de modelos mentales mecanicistas (causales) de de Kleer y Brown, así como de los propuestos por Williams, Hollan y Stevens. Por el momento es importante darse cuenta de que la concepción de modelo mental de Johnson-Laird es mucho más abarcadora que la de estos autores. (Eso deberá quedar claro en la sección siguiente).

En un artículo reciente, Ibrahim Halloun (1996), centrándose en lo que él llama la *modelización esquemática*, también se refiere a modelos mentales (p. 1021) :

“La modelización esquemática es una teoría epistemológica en desarrollo, fundamentada en la investigación cognitiva. Esta teoría sustenta que los modelos son componentes principales del conocimiento de cualquier persona y que la modelización es un proceso cognitivo básico para construir y usar el conocimiento en el mundo real. Tres de las más fundamentales premisas de la modelización esquemática son :

- 1.- *Construimos modelos mentales* que representan aspectos significativos de nuestro mundo físico y social, y manipulamos elementos de esos modelos cuando pensamos, planeamos e intentamos explicar eventos de ese mundo.
- 2.- Nuestra visión del mundo es causalmente dependiente tanto de cómo es el mundo como de cómo somos nosotros. Deriva de lo anterior un obvio pero importante corolario : “Todo nuestro conocimiento sobre el mundo depende de nuestra habilidad para construir modelos del mismo” (Johnson-Laird, 1983).
- 3.- Los *modelos mentales* son internos a las mentes de las personas. Son *tácitos* y no pueden explorarse directamente. Pueden, sin embargo, investigarse indirectamente por la vía de los *modelos conceptuales* con los cuales las personas *se comunican* con otras verbalmente, simbólicamente o pictóricamente (y/o vía *modelos físicos*, que son artefactos materiales). Los modelos conceptuales usados en la vida diaria son frecuentemente subjetivos, idiosincráticos y no estructurados coherentemente”.

La concepción de modelo mental de Halloun es, por lo tanto, la de Johnson-Laird, pero lo que él destaca es el proceso de modelización, que se retomará en este trabajo en la sección correspondiente a las investigaciones sobre modelos mentales.

### **La tipología de los modelos mentales de Johnson-Laird**

Teniendo en cuenta todos los principios (vínculos), referidos anteriormente, que restringen la naturaleza de los modelos mentales, Johnson-Laird propone lo que él llama (1983, p. 422) una tipología informal y tentativa para los modelos mentales.

En primer lugar, distingue entre *modelos físicos* que son los que representan el mundo físico y *modelos conceptuales*<sup>3</sup> que son los que representan cosas más abstractas. Después identifica seis tipos principales de modelos físicos :

---

<sup>3</sup> Los modelos conceptuales de Johnson-Laird no son los mismos de Norman (1983, p. 12) que son modelos precisos, consistentes y completos inventados por los profesores, investigadores, ingenieros y proyectados como instrumentos para la comprensión y la enseñanza. Los de Johnson-Laird son modelos que las personas tienen en sus cabezas y que representan estados de cosas abstractos en relación con los estados de cosas físicos, representados por los modelos físicos.

1.- *Modelo relacional* es un cuadro (“frame”) estático que consta de un número finito de elementos (“tokens”), que representan un conjunto finito de entidades físicas, de un conjunto finito de propiedades de los elementos, que representan propiedades físicas de las entidades, y de un conjunto finito de relaciones entre los elementos que representan relaciones físicas entre las entidades (ibid.).

2.- *Modelo espacial* es un modelo relacional en el que las únicas relaciones que existen entre las entidades físicas representadas son espaciales y el modelo representa estas relaciones localizando los elementos (“tokens”) en un espacio dimensional (típicamente de dos o tres dimensiones). Este tipo de modelo puede satisfacer las propiedades del espacio métrico ordinario, en particular la continuidad psicológica de sus dimensiones y la desigualdad triangular (la distancia entre dos puntos nunca es más que la suma de la distancia entre cada uno de ellos y un tercer punto cualquiera) (ibid.).

3.- *Modelo temporal* es el que consta de una secuencia de cuadros “frames” espaciales (de una determinada dimensionalidad) que se produce en un orden temporal que corresponde al orden de los eventos (aunque no necesariamente en tiempo real) (ibid.).

4.- *Modelo cinemático* es un modelo temporal que es psicológicamente continuo ; es un modelo que representa cambios y movimientos de las entidades representadas sin discontinuidades temporales. Naturalmente, este modelo puede funcionar (“rodar”) en tiempo real y ciertamente lo hará si fuera construido por la percepción (p. 423).

5.- *Modelo dinámico* es un modelo cinemático en el que existen también relaciones entre ciertos cuadros (“frames”) que representan relaciones causales entre los eventos representados (ibid.).

6.- *Imagen* es una representación, centrada en el observador, de las características visibles de un modelo espacial tridimensional o cinemático subyacente. Corresponde, por lo tanto, a una visión (o proyección) del objeto o evento representado en el modelo subyacente (ibid.).

Johnson-Laird dice que no hay una línea divisoria precisa entre percepción y concepción, pero encuentra plausible coincidir con Marr (1982, apud Johnson-Laird, 1983, p. 423) y suponer que la percepción normalmente produce modelos dinámicos, métricos, tridimensionales de estados de cosas del mundo, en los cuales cada cuadro (“frame”) caracteriza las formas volumétricas de los objetos y las relaciones espaciales entre ellos en términos de un sistema de coordenadas referido a los objetos. El único problema es la causalidad (por ser una relación abstracta), pero el sistema perceptivo parece ser sensible a ella, o mejor, a pistas de ella.

Entonces, estos seis modelos están clasificados por él como físicos en el sentido de que, con excepción de la causalidad, corresponden directamente al mundo físico. Pueden representar situaciones perceptibles, pero no relaciones abstractas o cualquier cosa además de las descripciones de situaciones físicas determinadas.

Los modelos mentales no derivados de la percepción pueden construirse para representar situaciones verdaderas, posibles o imaginarias. Tales modelos pueden, en principio, ser físicos o conceptuales, pero, en general, se construyen a partir del discurso y éste requiere un modelo conceptual. Los modelos conceptuales, por no tener el referente del mundo físico, exigen, más que los modelos físicos, un mecanismo de auto-revisión recursiva.

Johnson-Laird (p. 425) distingue cuatro tipos principales de modelos conceptuales<sup>4</sup>

<sup>4</sup> Se reitera que no son los mismos de Norman.

1.- *Modelo monádico* es el que representa afirmaciones (como aquéllas del razonamiento silogístico) sobre individualidades, sus propiedades e identidades. Este modelo tiene tres componentes : un número finito de elementos (“tokens”) que representan entidades individuales y sus propiedades ; dos relaciones binarias -identidad (=) y no identidad ( $\neq$ ) ; y una anotación especial para indicar que es incierto si existen determinadas identidades (p.425).

Por ejemplo, el modelo conceptual monádico de la aserción “Todos los licenciados son profesores” puede ser el siguiente :

$$\text{Licenciado} = \text{profesor} \\ (\text{profesor})$$

en donde el elemento (“token”) licenciado es una anotación<sup>5</sup> para indicar que el correspondiente elemento (“token”) mental representa un individuo que es licenciado. La anotación de incluir un elemento entre paréntesis corresponde a esa anotación especial de los modelos mentales conceptuales que indica que es incierto si la individualidad correspondiente existe o no en el dominio del modelo. Por tanto, puede haber un profesor no licenciado.

Los modelos monádicos pueden acomodar sólo aserciones simples de un único predicado que indica propiedades, identidades y no identidades. Para aserciones más generales es necesario emplear otro tipo de modelo, el relacional (ibid.).

2.- *Modelo relacional* es aquel que añade un número finito de relaciones, posiblemente abstractas, entre las entidades individuales representadas en un modelo monádico. Este modelo es necesario para representar una aserción como “existen más as que bes”, que requiere una representación del siguiente tipo (p. 425) :

$$a - b \\ a - b \\ a$$

3.- *Modelo metalingüístico* es el que contiene elementos (“tokens”) correspondientes a ciertas expresiones lingüísticas y ciertas relaciones abstractas entre ellas y elementos del modelo (de cualquier tipo, incluyendo el propio modelo meta-lingüístico) (ibid.).

Por ejemplo, la aserción “Uno de los operarios se llama Juan” requiere un modelo metalingüístico de la siguiente especie :

$$\text{“Juan”} \longrightarrow \begin{array}{l} \text{operario} \\ \text{operario} \\ \text{operario} \end{array}$$

en donde las comillas se están usando para significar una expresión lingüística y la flecha denota referencia : la expresión lingüística “Juan” se refiere ( $\rightarrow$ ) a tal operario.

4.- *Modelo conjunto teórico* es aquel que contiene un número finito de elementos (“tokens”) que representan directamente conjuntos ; puede contener también un conjunto finito de elementos (“tokens”) que representan propiedades abstractas del conjunto y un número finito de relaciones (incluyendo identidad y no-identidad) entre los elementos que representan conjuntos.

---

<sup>5</sup> Obviamente, estas anotaciones son arbitrarias ; es improbable que se vaya a saber cómo se representan estos elementos de hecho en la mente.

Por ejemplo, consideremos la aserción “Algunas bibliografías se relacionan a sí mismas y otras no” ; un modelo mental de la forma

$$b_1 = \begin{cases} \text{“}b_1\text{”} \\ \text{“}b_2\text{”} \\ \text{“}b_3\text{”} \end{cases} \quad (\text{p. 428})$$

representa el hecho de que una bibliografía  $b_1$  consta de tres nombres, incluyendo el suyo propio.

Análogamente, un modelo de la forma

$$b_6 = \begin{cases} \text{“}b_3\text{”} \\ \text{“}b_4\text{”} \\ \text{“}b_5\text{”} \end{cases} \quad (\text{ibid.})$$

representa una bibliografía de las bibliografías que no se relacionan a sí mismas.

En estos modelos, los elementos (“tokens”)  $b_1, b_2, \dots, b_6$  representan conjuntos (bibliografías) directamente y la llave representa la relación de inclusión.

Johnson-Laird califica esta tipología como informal y tentativa, pues, en última instancia, es la investigación la que va a decir cómo son los modelos mentales que las personas tienen en la cabeza. No obstante, los diferentes tipos de modelos físicos y conceptuales por él propuestos revelan el carácter esencial de los modelos mentales : derivan de un número relativamente pequeño de elementos y de operaciones recursivas sobre tales elementos ; su poder representacional depende de procedimientos adicionales para construirlos y evaluarlos ; las mayores restricciones sobre los mismos derivan de la estructura percibida o concebida de los estados de cosas del mundo, de los conceptos que subyacen a los significados de los objetos y eventos y de la necesidad de mantenerlos libres de contradicciones (p. 430).

### **La metodología de la investigación en modelos mentales**

Si una persona es capaz de andar por la casa a oscuras sin grandes problemas es porque tiene un modelo mental espacial de su casa. Si una persona es capaz de explicar cómo funciona una nevera es porque tiene un modelo mental relacional, o dinámico, de ese dispositivo físico. Si fuera capaz de dar significado a una aserción como “No todos los doctores en Física son investigadores” es porque tiene un modelo mental conceptual (en el sentido de Johnson-Laird) en el que existen elementos (“tokens”) que corresponden a “doctores en Física”, a “investigadores” y a la relación “no todos” que le permiten representar tal situación.

Se insiste en que los modelos mentales están en la cabeza de las personas. ¿Cómo, entonces, investigarlos ? ¿Cómo construir modelos mentales de los modelos mentales de las personas ? ¿ O modelos conceptuales (en el sentido de Norman) de los modelos mentales compartidos por varias personas (si es que eso existe) ?.

Las posibles metodologías para investigar los modelos mentales están basadas en la premisa de que las representaciones mentales de las personas pueden inferirse (modelarse) a partir de sus comportamientos y verbalizaciones. Dicho esto, se supone también que esos modelos pueden simularse en ordenador.

Todavía, sean cuales fueran esas metodologías, la investigación en esta área es bastante difícil, por dos razones fundamentales.

En primer lugar, porque no se puede preguntar simplemente a la persona cuál es el modelo mental que tiene para determinado estado de cosas, pues puede no tener plena conciencia de ese modelo. Puede decir que cree en alguna cosa y proceder en desacuerdo con esta creencia ; es decir, ella, de hecho, cree que cree, pero lo que hace no confirma eso. Las creencias de las personas, particularmente cuando son de naturaleza procedimental, no están disponibles para examen (Norman, 1983, p. 11). Además, cuando se pregunta a una persona por qué o cómo hizo alguna cosa, puede sentirse impelida a dar una razón, incluso aunque no la tuviese antes de que se haga la pregunta. Puede (usando un modelo mental de las expectativas de quien hace la pregunta) dar una razón que le parece que es la que la persona que hace la pregunta querría oír (ibid.). De ahí puede pasar a creer en esta razón a pesar de haberse generado en el momento de responder a la pregunta. (Por este motivo, los protocolos verbales que describen lo que la persona hace a medida que resuelve un problema, por ejemplo, son más confiables que las explicaciones).

En segundo lugar, porque no adelanta nada buscar modelos mentales claros, nítidos, elegantes, pues los modelos que las personas de hecho tienen son estructuras confusas, mal hechas, incompletas, difusas (Norman, 1983, p.14). Con este tipo de representación mental es con el que el investigador en esa área tiene que lidiar e intentar entender.

A pesar de las dificultades, el *análisis de los protocolos*, el uso de informaciones verbales del sujeto como fuente de datos ha sido, probablemente, la técnica más usada para investigar la cognición humana. No se trata, sin embargo, de introspección, pues en ésta las verbalizaciones del sujeto ya se constituyen en teoría acerca de sus propios procesos de pensamiento mientras que en el análisis de protocolos, las verbalizaciones se consideran como datos que deben ser explicados por teorías generadas por el investigador, tal vez en la forma de simulación computacional (Simon y Kaplan, 1989, p. 21).

Hay varias maneras de hacer que las personas generen protocolos verbales, incluyendo entrevistarlas, pedirles que hablen libremente, que piensen en voz alta, que describan lo que están haciendo cuando ejecutan una tarea. Los protocolos que se generan cuando la persona piensa en voz alta *en cuanto* desempeña una tarea como, por ejemplo, la solución de un problema, se llaman *protocolos concurrentes*. Cuando se pide al sujeto que diga todo lo que consigue recordar sobre la solución de un problema inmediatamente *después* de haberla obtenido, el protocolo recibe el nombre de *retrospectivo* (ibid.). (En este caso, hay que tener precaución, pues el sujeto puede ser capaz de reconstruir eventos que no ocurrieron).

Típicamente, las verbalizaciones que generan los protocolos se graban, se transcriben y se analizan a la luz de alguna teoría.

Los estudios relatados en la sección siguiente deberán ayudar a aclarar cómo el análisis de protocolos y otras técnicas han sido usadas en la investigación sobre modelos mentales.

### **Algunas investigaciones sobre modelos mentales**

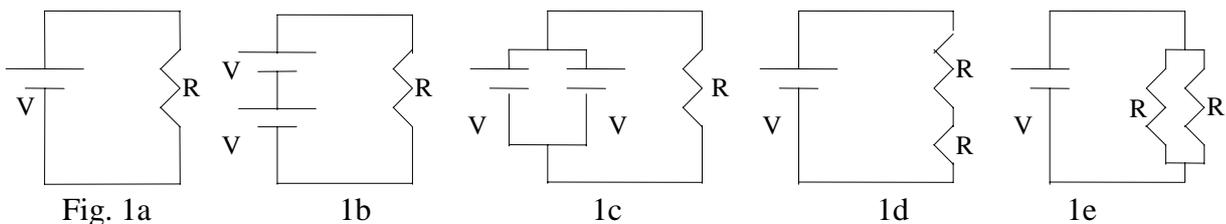
**Gentner y Gentner (1983)**, describen investigaciones en las que hicieron previsiones sobre la actuación de alumnos en problemas de circuitos eléctricos a partir de dos analogías que, según ellos, son los modelos más comúnmente usados por los estudiantes en esta área ; el modelo del “fluido en movimiento” (analogía hidráulica) y el modelo de la “multitud en movimiento”. En el primero, el fluido en movimiento corresponde a la corriente eléctrica, los conductos a los hilos, los

estrechamientos de los conductos a los resistores, los depósitos a las baterías y la diferencia de presión (función de la altura del agua en el depósito) a la diferencia de potencial. En el segundo, la cantidad de personas que pasa por un "portón" por unidad de tiempo corresponde a la corriente eléctrica, el portón (paso, salida) a la resistencia eléctrica y la "fuerza con la que las personas se empujan" a la diferencia de potencial; en este modelo no hay un análogo adecuado para las baterías.

En la previsión de estos autores, la adopción de uno u otro de estos modelos resultaría en actuaciones diferentes en problemas de circuitos eléctricos. Por ejemplo, el primer modelo podría llevar a conclusiones equivocadas sobre circuitos con resistores en serie o en paralelo: los sujetos podrían tender a considerar que, independientemente de cómo estén ligados, cuanto más resistores, mayor será la resistencia del circuito y menor la corriente. Por otro lado, el segundo modelo tendería a conducir a la previsión de que dependiendo de cómo estén dispuestos los "portones", la corriente será mayor o menor; por ejemplo, si los resistores estuvieran en paralelo (portones uno al lado del otro) la corriente será mayor (más gente pasará). Pero este modelo tendría dificultades con problemas que implicasen baterías en serie o en paralelo.

A partir de esa previsión hicieron un estudio con 36 estudiantes de escuela secundaria y novatos universitarios que poco sabían de electricidad (op. cit. p. 117). Dieron a los alumnos un folleto que contenía varias cuestiones sobre electricidad que debían responder, trabajando a ritmo propio. En la primera página había un circuito simple con una batería y un resistor, como el de la figura 1a. En las cuatro páginas siguientes había, en cada una de ellas, un circuito un poco más complicado como los de las figuras 1b, 1c, 1d y 1e.

En cada uno de estos casos, los alumnos debían decir si la corriente en el circuito era mayor, igual o menor que la del circuito más simple (1a). Después de dar sus respuestas a todas estas preguntas, debían describir, con sus propias palabras, cómo pensaban la electricidad en un circuito. A continuación, en la página siguiente del folleto, debían ser más específicos y contestar, en cada uno de los casos (1b, 1c, 1d y 1e) si habían pensado en un fluido escurriendo en objetos en movimiento o alguna otra visión de electricidad a medida que resolvían los problemas propuestos.



Los investigadores, entonces, analizaron los protocolos obtenidos a partir de las respuestas dadas e identificaron siete alumnos que usaron, consistentemente, en todas las cuestiones el modelo de fluido en movimiento y ocho que usaron el modelo de la multitud (objetos) en movimiento. Las respuestas de los alumnos que fueron inconsistentes en el uso de modelos no se consideraron en esta etapa de la investigación.

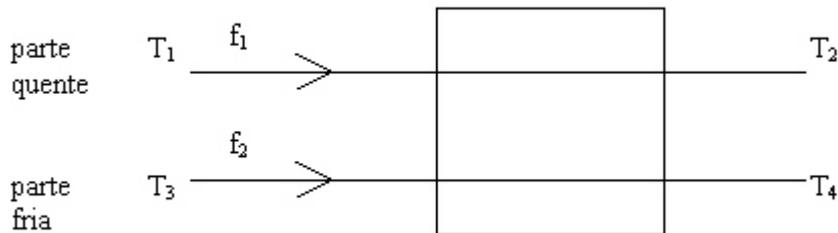
Los resultados obtenidos confirmaron la previsión de que los alumnos que usaron el modelo del fluido en movimiento se defendían mejor en las cuestiones sobre baterías que sobre resistores, mientras que los que prefirieron el modelo de la multitud en movimiento tenían mejor actuación en las cuestiones sobre resistores, particularmente en paralelo, que sobre baterías.

Como se puede desprender de esta descripción, el concepto de modelo mental de Gentner y Gentner es, prácticamente, el mismo de analogía, en el sentido más tradicional -- aquel en el que el análogo guarda una correspondencia muy próxima con aquello que representa. La definición de Johnson-Laird -- modelo mental como análogo estructural de un estado de cosas (un evento o un objeto) del mundo -- parece ser más abarcadora.

**Williams, Hollan y Stevens**, aquellos autores que definen modelo mental como un conjunto de objetos autónomos<sup>6</sup> “conectados” (1983, p. 133), relatan experimentos (op. cit. p. 135-146) sobre modelos mentales de un sistema de refrigeración que ellos denominan “cambiador de calor” (“heat exchanger”). Tal sistema consiste en lo siguiente (p. 132) :

“La función de un “cambiador de calor” es enfriar un fluido caliente. Este fluido puede ser el agua o el aceite usado para lubricar y enfriar una máquina. El calor es retirado (sic) por medio de un fluido frío, en general agua de río o de mar. Los parámetros importantes del funcionamiento del “cambiador de calor” son el flujo de fluido caliente ( $f_1$ ), las temperaturas de entrada y de salida del fluido caliente ( $T_1$  y  $T_2$ ), el flujo de fluido frío ( $f_2$ ) y las temperaturas de entrada y salida del flujo frío ( $T_3$  y  $T_4$ )”.

Esquemáticamente, el cambiador de calor puede representarse de la siguiente forma (ibid.) :



Para estos autores, el comportamiento (i.e., cambios en los parámetros) de un objeto autónomo es gobernado por un conjunto de reglas internas. En el caso del “cambiador de calor”, sugieren que un conjunto de ocho reglas sería suficiente para dar cuenta de su funcionamiento (op. cit. p. 136) :

$$\begin{array}{lcl}
 dT_1 & \Rightarrow & dT_2 \\
 dT_1 & \Rightarrow & dT_4 \\
 df_1 & \Rightarrow & dT_2 \\
 df_1 & \Rightarrow & dT_4 \\
 dT_3 & \Rightarrow & dT_2 \\
 dT_3 & \Rightarrow & dT_4 \\
 df_2 - & \Rightarrow & dT_2 \\
 df_2 - & \Rightarrow & dT_4
 \end{array}$$

<sup>6</sup> Objeto autónomo es, para Williams, Hollan y Stevens (1983), un objeto mental que representa explícitamente alguna cosa, cuyas conexiones topológicas con otros objetos son también explícitas, y que tiene un cierto número de parámetros internos.

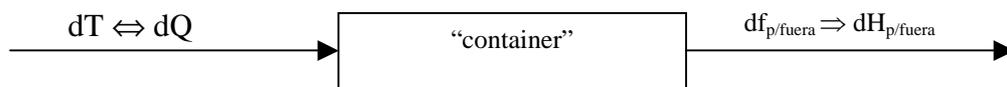
El símbolo  $\Rightarrow$  significa que cualquier cambio en el parámetro de la izquierda causa un cambio en la misma dirección en el parámetro de la derecha. El símbolo  $-\Rightarrow$  significa que un cambio en el parámetro de la izquierda causa un cambio en la dirección opuesta en el parámetro de la derecha.

Éste es un modelo “compilado” de “cambiador de calor” que un conocedor de máquinas térmicas podría tener sobre el funcionamiento superficial de ese dispositivo. Superficial porque estas reglas no son suficientes para representar un entendimiento profundo de los mecanismos subyacentes a un sistema térmico como éste.

Este conjunto de reglas puede también interpretarse como el que define el funcionamiento de un objeto autónomo aislado, que sería, entonces, un caso degenerado de modelo mental (i.e., un modelo mental constituido por un solo objeto autónomo).

En los experimentos que hicieron, esos investigadores pedían a los sujetos que pensasen en voz alta al tiempo que respondían a una serie de cuestiones sobre los valores de los parámetros  $T_1, T_2, T_3, T_4, f_1$  y  $f_2$  y sobre los efectos cualitativos de variaciones en esos parámetros. Las respuestas de los sujetos y lo que ellos decían mientras pensaban en voz alta generaron los protocolos cuyo análisis permitió identificar tres modelos de “cambiador de calor” (p. 137-146) :

En el modelo 1 - Modelo de “Container” - el sistema se representa como un “container” hacia dentro del cual se está bombeando calor y la manera de salir de él es a través del flujo de fluido, agua o aceite, hacia fuera. Esquemáticamente, ese modelo sería así :



La regla  $dT \Leftrightarrow dQ$  dice que una variación en la temperatura ( $T$ ) da como resultado una variación equivalente en el calor total ( $Q$ ) del fluido y viceversa. La regla  $df_{p/fuera} \Rightarrow dH_{p/fuera}$  dice que un aumento en el flujo de fluido hacia fuera implica un aumento en el flujo de calor hacia fuera del sistema.

Este modelo da cuenta de las cuatro primeras reglas del “modelo del especialista”, pero es reticente en relación con las cuatro últimas, aunque no viole ninguna de ellas (p. 138). En la concepción de los investigadores, el modelo está constituido por cuatro objetos autónomos (la entrada de calor, el “container” y las dos salidas, agua o aceite) y tres puertos (conexiones mentales entre objetos autónomos en un modelo mental).

Los otros dos modelos son suplementarios al modelo 1, conteniendo más objetos autónomos y puertos necesarios para justificar respuestas a determinadas cuestiones.

El concepto de modelo mental usado en esta investigación, semejante al de modelo mecanicista propuesto por de Kleer y Brown (1983, p. 155), es también más restringido que el de Johnson-Laird. Además, se trata de un modelo básicamente proposicional.

**Gutiérrez y Ogborn (1992)** usaron el concepto de modelo mental mecanicista propuesto por de Kleer y Brown (1983) para analizar protocolos relativos a fuerza y movimiento, tanto de los sujetos de su investigación como de otros estudios ya publicados por otros autores. El modelo de de Kleer y Brown ya se ha descrito en este trabajo (p. 14) ; sin embargo, vale la pena retomarlo en la interpretación de Gutiérrez y Ogborn (op. cit. p. 201-203), con el fin de aclarar aún más este concepto de modelo mental.

Según estos autores, el modelo mecanicista de de Kleer y Brown intenta responder a la siguiente pregunta : ¿ qué necesita un sistema cognitivo que se encuentra con un sistema físico (como un dispositivo hidráulico, eléctrico o térmico) para ir desde cómo está hecho hasta una o más posibilidades de cómo funciona, suficientemente buenas para explicar lo que hace ? (ibid. p. 201).

Este proceso puede analizarse en cuatro etapas :

- representar el sistema ;
- atinar con un modelo de cómo podría funcionar ;
- imaginar el modelo funcionando (simulación mental) ;
- comparar con la realidad los resultados imaginados del modelo.

Estas etapas se repiten si la última no es satisfactoria. Cinco nociones básicas introducidas por de Kleer y Brown están implícitas en estas etapas :

<i>topología del dispositivo</i> :	una representación de la estructura del sistema físico ;
<i>visionamiento</i> (“envisioning”) :	ir desde la estructura hasta cómo podría funcionar el sistema ;
<i>modelo causal</i> :	resultado del proceso de visionamiento ;
<i>ejecución</i> (rodaje, “running”) :	imaginar lo que haría el modelo causal ;
<i>episodio</i> :	intervalo de tiempo durante el que la explicación permanece igual.

El modelo de de Kleer y Brown atribuye al sistema cognitivo un compromiso ontológico básico : todo tiene una causa ; el razonamiento con este modelo es causal, no legal ; las explicaciones son en términos de acciones y efectos, no de leyes y regularidades (ibid. p. 202). El sistema cognitivo puede incluso generar, i.e., inventar causas. Las causas que se inventan sólo porque son necesarias reciben el nombre de míticas.

Este tipo de modelo es básicamente cualitativo. En él no existen leyes y relaciones cuantitativas, solamente relaciones del tipo “si esto ocurre, entonces tal cosa se producirá”, usando propiedades tal como “grande”, “pequeño”, “mayor”, “menor”, “igual”, “mismo”.

La topología, o visionamiento, el modelo causal y la ejecución pueden discriminarse como sigue (ibid.):

*Topología* - materiales : partes cuyos atributos pueden verse afectados por una acción causal.  
 componentes : partes que pueden efectuar un cambio causal.  
 conductos : partes que pueden conducir materiales o acciones causales.

*Visionamiento* - causas cuyos efectos están ausentes, o son muy pequeños, son ignoradas ; hasta probar lo contrario, los atributos desconocidos tiene valores despreciables.

*Modelo causal* - *principio de la localidad* : la causa está estructuralmente cerca de su efecto. Si es la estructura, como un todo, lo que determina la causa, entonces es no local.  
*principio de la asimetría* : las causas preceden a sus efectos. Según este principio, los episodios se construyen en una secuencia dada determinada por relaciones causales.

Estos dos principios son de de Kleer y Brown ; sin embargo Gutiérrez y Ogborn juzgan necesario añadir otros tres :

*principio de la productividad* : si existe un efecto, habrá sido siempre producido por una causa ;

*principio de la constancia* : si existe una causa, invariablemente habrá un efecto ;  
*principio de la unicidad* : la misma causa siempre producirá el mismo efecto.

*Ejecución* - un modelo aceptable, cuando es imaginado en funcionamiento, debe tener consistencia, correspondencia y robustez. Consistencia significa que el modelo no debe tener contradicciones internas ; correspondencia requiere que prediga lo que de hecho ocurre; robustez implica que el modelo siga haciendo previsiones correctas cuando el contexto cambia a otro similar.

Si estos vínculos no fueran satisfechos, el modelo debe reconstruirse (reformularse) : una nueva causa puede introducirse o un nuevo efecto, o la topología debe modificarse. Esta reconstrucción recibe el nombre de “aprendizaje” por de Klee y Brown (apud. Gutiérrez y Ogborn, p. 203), no obstante, no implica “enseñanza”, sólo observaciones y reflexión.

Como se dijo anteriormente, Gutiérrez y Ogborn utilizaron el modelo de de Klee y Brown para analizar protocolos relativos a fuerza y movimiento. Trabajaron con 10 estudiantes españoles, cinco de la edad de 13-14 años que habían tenido apenas una introducción elemental a la mecánica de Newton en la disciplina de ciencias y cinco en la de 17-18 años, que habían cursado ya dos años de mecánica newtoniana.

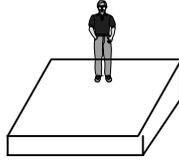
La técnica utilizada fue la de entrevista clínica, proponiendo a los estudiantes situaciones sacadas de historias y/o viñetas que implican movimientos cotidianos. Con eso pretendían apartarse de las situaciones problemáticas normalmente propuestas en el aula y dejar a los alumnos más a voluntad. Las entrevistas duraron aproximadamente una hora, fueron grabadas en cinta magnética y transcritas.

En el análisis de los protocolos así obtenidos, las explicaciones de los movimientos dadas por los alumnos se dividieron en episodios ; cada vez que cambiaba la naturaleza de la causa, se consideraba un nuevo episodio.

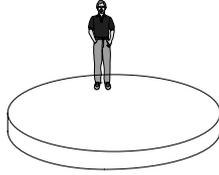
De su investigación y del reanálisis de los protocolos obtenidos por otros investigadores, Gutiérrez y Ogborn concluyeron que, de modo general, el modelo de de Klee y Brown es adecuado para describir el razonamiento causal en una amplia variedad de situaciones. En particular, observaron que frecuentemente los entrevistados cambiaban el modelo causal, modificando la naturaleza de la explicación en uno o más episodios y eso es una cosa de la que el modelo da cuenta (el cambio de idea, o aprendizaje, de de Klee y Brown). Encontraron también evidencias de la causalidad mítica y de los varios principios del modelo causal.

**Stella Vosniadou (1994)** dice que en sus investigaciones sobre cambio conceptual ha conseguido identificar un número relativamente pequeño de modelos mentales, concernientes a los conceptos tratados en una entrevista, que los estudiantes usan consistentemente. Por ejemplo, en relación con el concepto de Tierra, esta autora encontró que el 80% de los niños entrevistados utilizaban de manera consistente uno de los siguientes modelos (p. 53) :

1. Rectangular



2. Disco



3. Dual



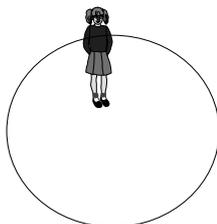
4. Esfera hueca



5. Esfera achatada



6. Esfera



Los dos primeros se han denominado “iniciales” porque parecen estar basados en la experiencia cotidiana y no presentan influencia del modelo científico, culturalmente aceptado, de Tierra esférica. Los tres siguientes se han considerado sintéticos porque combinan aspectos de los modelos iniciales y del modelo científico : el modelo dual, según el cual hay dos Tierras, una plana en la que viven las personas, y otra esférica que es un planeta en el cielo; el de la esfera hueca en la que viven las personas en una región plana interna; el de la esfera achatada según el cual las personas viven en regiones planas encima o debajo (ibid.).

Vosniadou interpreta el cambio conceptual como una modificación progresiva de los modelos mentales que la persona tiene sobre el mundo físico, conseguido por medio de *enriquecimiento* o *revisión*. Enriquecimiento implica la adición de informaciones a las estructuras conceptuales existentes ; revisión supone cambios en las creencias o presupuestos individuales o en la estructura relacional de una teoría.

Esta autora establece una distinción entre una *teoría física estructural ingenua* (“naive framework theory of physics”) que se constituye muy temprano en la infancia y *teorías específicas* que describen la estructura interna de dominios conceptuales en los que los conceptos están “encajados”. Cuestiona la hipótesis -- que es prácticamente de sentido común -- de que la construcción del conocimiento empieza por la formación de conceptos aislados que van relacionándose hasta que constituyen estructuras conceptuales complejas. Su hipótesis es la de que los conceptos están “encajados”, desde el principio, en estructuras teóricas más amplias. Argumenta, por ejemplo, que, desde muy pequeños, los niños desarrollan cinco vínculos sobre el comportamiento de objetos del mundo físico -- continuidad, solidez, no-acción-a-distancia, gravedad e inercia -y que estas “presuposiciones atrincheradas” (op. cit. p. 47) están organizadas en una *teoría física estructural ingenua* no disponible para verificación consciente. Esta teoría restringe el proceso de adquisición de conocimiento sobre el mundo físico.

Las teorías específicas constan de un conjunto de proposiciones o creencias interrelacionadas que describen las propiedades y comportamientos de determinados objetos físicos.

El cambio conceptual por revisión puede ocurrir tanto en las teorías específicas como en la teoría estructural, pero en este caso es mucho más difícil y es probable que genere las llamadas concepciones alternativas (“misconceptions”). Esta investigadora interpreta las concepciones alternativas como intentos del aprendiz para interpretar la información científica a la luz de una teoría estructural que contiene informaciones contradictorias con la visión científica.

En el caso de la Tierra, los niños tienen dificultad para construir mentalmente el modelo científico porque este modelo viola ciertas “presuposiciones atrincheradas” de la teoría física estructural ingenua en la que el concepto de Tierra esta “encajado”.

El concepto de modelo mental de Vosniadou es el de “representación analógica que generan los individuos durante el funcionamiento cognitivo y que tiene la característica especial de preservar la estructura de aquello que supuestamente representa” (op. cit. p. 48):

“Los modelos mentales son representaciones dinámicas y generativas que pueden manipularse mentalmente para proveer explicaciones causales de fenómenos físicos y hacer previsiones sobre estados de cosas del mundo físico. Se supone que muchos modelos mentales se crean en el momento para resolver cuestiones de situaciones problemáticas específicas. Sin embargo, es posible que algunos modelos mentales, o parte de ellos, que fueron útiles una vez, se almacenen como estructuras separadas y se recuperen de la memoria de largo plazo cuando sea necesario”.

Esta definición es prácticamente la de Johnson-Laird. A partir de ella, Vosniadou realizó varias investigaciones sobre cambio conceptual tratando el concepto de Tierra, explicaciones sobre el día y la noche, el concepto de fuerza y el de calor. En todos los casos, encontró que los sujetos usan consistentemente uno de un pequeño conjunto de modelos mentales.

La metodología utilizada en estos estudios consiste en formular varias cuestiones sobre el concepto que se está investigando. Algunas de ellas requieren una respuesta verbal, otras estimulan la realización de diseños y otras implican la construcción de modelos físicos. La suposición que está por detrás es la de que los estudiantes “acceden” al conocimiento relevante y construyen un modelo mental que les permite responder a las cuestiones propuestas. La investigadora, entonces, busca entender y describir esos modelos y usarlos para hacer inferencias sobre la naturaleza de las teorías específicas y estructurales que los restringen.

Se destacan dos aspectos de esa metodología (p. 50) : el *tipo de cuestiones* utilizadas y el *test de consistencia interna*.

Las cuestiones son “*generativas*” (productivas, no factuales), i.e., cuestiones que no pueden responderse a través de simples repeticiones de información no asimilada. En el caso de la Tierra, preguntas como : ¿cuál es la forma de la Tierra ? o ¿la Tierra se mueve ? no sirven porque respuestas correctas no significan necesariamente que los alumnos hayan entendido el concepto en cuestión. Preguntas como “Si usted caminase muchos días siempre en línea recta ¿a dónde llegaría ?” o “¿La Tierra tiene fin ? tendrían mayor potencial para hacer que los alumnos buscasen en su base de conocimiento la información relevante para construir un modelo mental de la Tierra o recuperasen de la memoria de largo plazo un modelo ya construido.

El test de consistencia interna consiste en verificar si el patrón de respuestas de un sujeto dado para todas la preguntas “generativas” relativas al concepto investigado puede explicarse por el uso consistente de un único modelo mental genérico subyacente.

**Harrison y Treagust (1996).** Hicieron un estudio sobre modelos mentales de 48 alumnos de octava a décima serie relativos a átomos y moléculas. Estos autores usan el término “modelo mental” para describir *sus* interpretaciones de las concepciones de átomos y moléculas de los alumnos. Estos alumnos eran de tres escuelas australianas diferentes y participaron voluntariamente en la investigación. Se utilizaron entrevistas semi-estructuradas con una duración media de 20 minutos.

Al principio de la entrevista, cada alumno recibía un trozo de lámina de aluminio y un bloque de hierro y se le preguntaba : “¿De qué crees que están hechos estos materiales ?”. Normalmente, el estudiante decía que el aluminio y el hierro estaban hechos de átomos y moléculas. Cuando esto no ocurría después de cuatro o cinco preguntas del tipo de la inicial, el investigador daba una pista usando el término “átomo”. A continuación, se pedía al alumno que pensase sobre su modelo mental (sic) de átomo y que lo dibujase en una hoja de papel y describiese el dibujo. (op. cit. p. 515).

Como la mayoría de los entrevistados diseñaba o mencionaba una bola o una esfera, se les daba una bola de poliestireno de 5 cm de diámetro y un pompón (con núcleo duro) y se les preguntaba cuál de esos dos modelos tenía alguna semejanza con su diseño y descripción (op. cit. p. 516).

A continuación, los alumnos recibían una hoja que contenía seis “diagramas de átomos”, sacados de libros didácticos y/o usados comúnmente por los profesores, y debían indicar cuál de

esos diagramas se ajustaba mejor a su modelo mental de átomo, cuál era el segundo mejor ajuste, cuál el tercero (si era posible) y cuáles eran los diagramas que no gustaban. En esta ocasión, frecuentemente los alumnos hablaban de núcleo, capa electrónica, nube electrónica, movimiento del electrón, protones y neutrones. Cuando esto no pasaba, el entrevistador daba pistas y preguntaba a los alumnos sobre nubes y capas electrónicas (ibid.).

La discusión pasaba, entonces, a moléculas y cada entrevistado recibía dos modelos concretos de molécula de agua, de los que debía escoger uno y explicar las razones de su elección.

Cada entrevista fue grabada en cinta cassette y transcrita literalmente. Las transcripciones y los dibujos de los alumnos formaron el conjunto de datos que iban a ser analizados. Del análisis cualitativo hecho, se identificaron categorías en las que podían encuadrarse las preferencias de los alumnos. En el artículo (Harrison y Treagust, 1996), se presentan tablas con tales preferencias y con los atributos atómicos y moleculares según los modelos de los estudiantes. Se presentan y comentan, también, varias descripciones de los alumnos sobre sus modelos mentales.

Obsérvese que la metodología usada en esta investigación probablemente no sería compartida por otros autores que también usan entrevistas, pues muchas cosas parecen sugerirse a los entrevistados. Los modelos se presentan al sujeto y él o ella debe escoger de entre ellos, el que más se asemeja al suyo. Este tipo de entrevista “no tan clínica” se ha usado en la investigación en enseñanza de las ciencias. La entrevista “teachback”, por ejemplo, consiste en una conversación entre entrevistador y entrevistado hasta que llegan a un consenso sobre el pensamiento del entrevistado (Pintó y col., 1996).

**Greca y Moreira (1996, 1997)** dirigieron una investigación con 50 estudiantes de ingeniería en una disciplina de Física General, en la que se propusieron investigar el tipo de representación mental usado por los alumnos cuando trabajaban con el concepto de campo, particularmente en el dominio del electromagnetismo, al resolver problemas y cuestiones conceptuales. El estudio se desarrolló en dos semestres consecutivos y tuvo como base conceptual la teoría de Johnson-Laird. Como se ha expuesto detalladamente en las primeras secciones de este trabajo, para Johnson-Laird existen tres clases de representaciones mentales : modelos mentales, proposiciones e imágenes. Los modelos mentales son análogos estructurales del mundo ; las proposiciones son cadenas de símbolos expresables en lenguaje natural e interpretadas (en términos de verdaderas o no) a la luz de modelos mentales ; las imágenes son visiones de los modelos mentales.

La investigación tuvo por objeto, entonces, identificar si los alumnos, al resolver problemas y cuestiones de Física, operaban mentalmente con modelos, proposiciones o imágenes, o con alguna combinación de estas representaciones mentales propuestas por Johnson-Laird.

Todo el estudio se realizó en condiciones normales de aula. En el primer semestre, con 25 alumnos, la metodología de enseñanza fue totalmente individualizada : el contenido se dividió en 20 unidades de estudio. Para cada unidad había una guía de estudio y un test escrito ; cuando se aprobaba el test, el alumno pasaba a la unidad siguiente. Había posibilidad de repetir el test tantas veces como fuera necesario para dominar el contenido de la unidad. El alumno trabajaba con ritmo propio y contaba con la asistencia de monitores y del profesor al preparar cada unidad de estudio. En esta metodología, si se desea, es posible alcanzar un elevado grado de interacción personal entre los estudiantes y el profesor, así como entre los estudiantes y los monitores. En este caso, los dos investigadores actuaron como profesores y monitores durante todo el semestre y estimularon al máximo posible la interacción personal ; procuraban discutir Física con los alumnos siempre que había oportunidad. Además, la evaluación de cada unidad era también dialogada.

A partir de esa fuerte interacción personal y del examen de las respuestas de los alumnos a las cuestiones propuestas en las guías de estudio y en los tests de evaluación, fue posible distinguir entre estudiantes que construyeron o no un modelo mental de trabajo para el concepto de campo electromagnético. En la primera categoría se encuadraron siete alumnos y sus modelos se clasificaron como modelos básicamente proposicionales (matemáticos) o esencialmente analógicos (imagísticos). En la categoría de los no modeladores se incluyeron 14 estudiantes. De un modo general, trabajaban con proposiciones sueltas (no articuladas en un modelo) y no usaban imágenes. Cuatro casos se consideraron intermedios entre las dos categorías.

En el segundo semestre de la investigación, con otros 25 alumnos, los investigadores continuaron con el mismo objetivo de identificar el tipo de representación mental usado por los alumnos; sin embargo, cambiaron la metodología de enseñanza y recogieron más materiales de análisis. El curso tenía tres periodos semanales de clases, de dos horas cada uno. En cada periodo, en un primer momento, había una parte expositiva de 30 a 35 minutos; después, los alumnos trabajaban en pequeños grupos (máximo cuatro participantes), generalmente resolviendo problemas y cuestiones de una lista. En cinco ocasiones, la lista se sustituyó por un experimento de laboratorio y en tres por un mapa conceptual. Al final de cada clase, los estudiantes entregaban el producto de su trabajo al profesor o al ayudante. Uno de los investigadores actuó como profesor y el otro como ayudante; los dos estuvieron siempre presentes y durante el trabajo en grupo interactuaron el máximo posible con los alumnos.

El curso tuvo tres evaluaciones formales individuales y una pequeña entrevista final también individual.

Todas las respuestas de los alumnos a las listas de problemas y cuestiones de cada clase o a las guías de laboratorio, todas las pruebas y todos los mapas conceptuales se utilizaron como material de análisis acoplados a las observaciones hechas durante la interacción personal y a la entrevista final.

Todo este material generó variables y puntuaciones: “concepto”, “problema”, “laboratorio”, “mapa conceptual”, “trabajo en grupo” y “modelo”. La variable “modelo” se construyó a partir de los indicadores obtenidos en el estudio del primer semestre, del análisis cualitativo del material escrito producido por los alumnos y, sobre todo, de la observación hecha por los investigadores a lo largo del semestre, interactuando bastante con los estudiantes. Esta variable recibió puntuaciones de 0 a 5 correspondientes a seis categorías:

- Categoría 0 (N=3).- estudiantes que operaban en la base del ensayo-y-error;
- Categoría 1 (N=3).- alumnos que trabajaban exclusivamente con proposiciones, no obstante sueltas, memorizadas mecánicamente;
- Categoría 2 (N=3).- alumnos que usaban básicamente proposiciones, pero con alguna articulación;
- Categorías 3 (N=7), 4 (N=7) y 5 (N=2).- estudiantes que construyeron algún modelo mental para el concepto de campo eléctrico.

A las demás variables también se le atribuyeron puntuaciones y se construyó una matriz de correlaciones. Se observó en la misma que las correlaciones fueron estadísticamente significativas y que las más bajas en relación con la variable “modelo” se dieron con las variables “laboratorio” y “trabajo en grupo”, mientras que las más altas se han mostrado con las variables “mapa conceptual” y “concepto”.

Tales correlaciones podrían, en principio, esperarse pues las variables “concepto” (construida a partir de las respuestas declarativas de los alumnos) y “mapa conceptual” eran las más

“conceptuales”, las que, a su vez, serían más dependientes del tipo de representación mental utilizado.

Un análisis factorial mostró que las variables “concepto”, “laboratorio”, “problema”, “mapa conceptual” y “trabajo en grupo” corresponden a un único factor que correlaciona 0,72 ( $p=0.0001$ ) con la variable “modelo”.

Estos resultados sugieren que la variable “modelo” explica buena parte de las puntuaciones de las demás variables, especialmente de las “conceptuales”.

Esta investigación fue, según los autores (Greca y Moreira, 1996, 1997), el comienzo de un programa<sup>7</sup> de investigación sobre modelos mentales y el aprendizaje de la Física que partió de cero en 1994. En este primer estudio, sólo intentaron, y aparentemente consiguieron, distinguir entre alumnos que trabajaban y no trabajaban con modelos mentales, según la teoría de Johnson-Laird, mientras ejecutaban tareas instruccionales de Física.

Desde el punto de vista metodológico, este estudio también usó el análisis cualitativo de documentos y verbalizaciones de los alumnos, pero, a diferencia de otros autores ya descritos, se completó con un tratamiento cuantitativo.

Los resultados obtenidos sugieren que en los cursos introductorios universitarios de Física, la mayoría de los alumnos trabaja con proposiciones no integradas o no interpretadas en un modelo mental. Las proposiciones que usan son definiciones y fórmulas manipuladas mecánicamente para resolver problemas o cuestiones. Algunos, sin embargo, dan evidencia de construcción de modelos y eso parece caracterizar un aprendizaje más significativo.

**Ibrahim Halloun (1996)** hizo una investigación sobre *modelización esquemática* que va en la misma línea de los resultados de Greca y Moreira, pues en la raíz de su investigación está el principio de que, en Física, el aprendizaje del alumno será tanto más significativo cuanto mayor sea su capacidad de modelar. Este autor argumenta que la Física es una ciencia de modelos y la modelización es una actividad sistemática de los físicos para construir y aplicar el conocimiento científico. Aprender Física supone, entonces, aprender a jugar el “juego de la modelización”.

Como se ha dicho en la página 16, Halloun usa el concepto de modelo mental de Johnson-Laird y piensa que, a pesar de ser tácitos, los modelos mentales pueden explorarse, indirectamente, a través de los modelos conceptuales que la persona externaliza de manera verbal, simbólica o pictórica al comunicarse con otra persona. Supone también que a través de instrucción adecuada, los modelos conceptuales -- que son generalmente subjetivos, idiosincráticos y no coherentemente estructurados -- pueden volverse relativamente objetivos y estructurados de manera coherente. Esta evolución, según él, se alcanza mejor en ciencias y en matemática en donde los modelos desempeñan un papel central.

En la idea de “instrucción adecuada” entra la modelización esquemática, o proceso de modelización tal como lo propone Halloun, cuyos cinco estadios son (op. cit. p. 1026-1028) :

- *selección* (identificación y descripción de la composición de cada sistema físico de la situación física y del respectivo fenómeno) ;
- *construcción* del modelo (conceptual o físico ; en el caso de la resolución de problemas, los alumnos son guiados en la construcción de un modelo conceptual matemático) ;
- *validación* del modelo (verificación de la consistencia interna) ;

---

<sup>7</sup> Modelos Mentales y Aprendizajes de Física según la Tipología de Johnson-Laird, apoyo CNPq.

- *análisis* del modelo (en el caso de la resolución de problemas, consiste principalmente en ejecutar el modelo matemático, obtener respuestas a los problemas e interpretarlas),
- *desdoblamiento* del modelo (tiene que ver con la transferencia).

Debe quedar claro, no obstante, que la modelización esquemática, como estrategia instruccional, se refiere principalmente a modelos conceptuales que el alumno externaliza. Los cinco estadios del proceso de la modelización sugeridos por este investigador no son jerárquicos; los tres estadios del medio se superponen; el primero y el segundo también pueden superponerse pues, en la práctica, en particular en la resolución de problemas, la construcción del modelo puede limitarse a la selección de un modelo matemático adecuado.

Halloun trabajó con 87 estudiantes libaneses, 59 de escuela secundaria y 28 universitarios, a los que les dio clases de recuperación durante el verano, pues no habían aprobado Física durante el semestre regular. A los estudiantes de secundaria les dio cinco clases de dos horas, en las que usó la modelización esquemática para enseñarles problemas de estática. A los universitarios les dio también cinco clases en las que les enseñó la modelización esquemática en mecánica de partículas.

Los problemas utilizados fueron de los normalmente encontrados en los libros de texto de mecánica. Se aplicaron pre y post-test y se compararon los resultados. Según el autor, tales resultados implican que la modelización esquemática es una estrategia válida para la enseñanza de la Física (op. cit. p. 1035). Sin embargo, su investigación tiene fallos metodológicos (por ejemplo, falta de grupo control) que limitan su validez interna y externa.

Se precisa más investigación sobre modelización conceptual (aquella que se puede enseñar) y modelización mental (aquella que se intenta investigar y, tal vez facilitar en la instrucción) y su interrelación. El valor de la investigación de Halloun es que parece intentar algo en este sentido.

**Lagrecia y Moreira (1998, 1999)** han realizado un estudio semejante al de Greca y Moreira (1996, 1997) en términos metodológicos, pero en el área de mecánica clásica introductoria, con el objetivo de identificar el tipo de representación mental utilizado predominantemente por los alumnos (estudiantes de ingeniería, en la asignatura Física I). Asimismo, buscaban también detectar “núcleos conceptuales”, es decir conceptos o grupos de conceptos que aparecerían con frecuencia a lo largo del desarrollo de la asignatura en las respuestas escritas de los alumnos, así como en sus mapas conceptuales y en la entrevista final. Dichos núcleos conceptuales podrían dar evidencias de posibles modelos mentales de los alumnos en mecánica. Esos investigadores han interactuado mucho con los estudiantes y han recogido y analizado una gran cantidad de representaciones externas (dibujos, diagramas, esquemas, resoluciones de problemas, mapas conceptuales, verbalizaciones) de ellos. El estudio fue conducido en el aula durante un semestre lectivo. En una categorización respecto a 13 alumnos, los investigadores encontraron, análogamente a Greca y Moreira, que un grupo de estudiantes ha trabajado básicamente con proposiciones aisladas no referidas a modelos mentales; un segundo grupo de alumnos que parecían trabajar con modelos mentales netamente proposicionales, o sea, reglas articuladas en modelos mentales y un tercer grupo de estudiantes que parecían operar con modelos principalmente imagísticos. Respecto a los núcleos conceptuales, han obtenido evidencias también de tres categorías principales: alumnos que tenían el concepto de fuerza relacionado directamente con el movimiento (fuerza como el agente que causa movimiento; todo movimiento necesita fuerza y toda fuerza involucra movimiento); estudiantes que parecían tener un modelo mental asociado a la tríada *velocidad constante*  $\Rightarrow$  *aceleración nula*  $\Rightarrow$  *fuerza nula* (este modelo les permitía hacer inferencias adecuadas en situaciones involucrando la primera y la segunda leyes de Newton); finalmente, una tercera categoría de alumnos que parecían tener un modelo mental de fuerza bastante próximo al científicamente aceptado, incluyendo, aparentemente, la idea de fuerza como interacción.

## Volviendo a la cuestión metodológica

Con excepción de la investigación de Halloun descrita en la sección anterior, que, en realidad, no investigó modelos mentales, todas las demás hicieron uso del análisis cualitativo de protocolos verbales y documentos (dibujos, esquemas, soluciones de problemas, mapas conceptuales, ...) producidos por los sujetos investigados en entrevistas o tareas instruccionales.

Eso es consistente con el hecho de que “los modelos mentales están en las cabezas de las personas” y la única manera de investigarlos es, indirectamente, a través de aquello que externalizan verbalmente, simbólicamente o pictóricamente.

Es interesante advertir que esa metodología es semejante a la que se desarrolló en la “Escuela de Würzburg”, en Alemania, al final del siglo pasado y principios de éste, por los psicólogos Oswald Külpe y Karl Bühler, denominada “introspección experimental sistemática”. Külpe fue discípulo de Wilhelm Wundt, considerado el fundador de la ciencia experimental de la psicología. Wundt instituyó en la Universidad de Leipzig, también en Alemania, en 1875, un laboratorio de psicología que se hizo famoso y atrajo a muchos estudiantes interesados en esa nueva ciencia. Entre éstos estaba Külpe que más tarde fue profesor en la Universidad de Würzburg y creó su propio laboratorio que, más adelante, comenzó a rivalizar con el de Wundt en importancia (Schultz y Schultz, 1995).

La “introspección experimental sistemática” de Külpe implicaba la realización de una tarea compleja (como el establecimiento de relaciones lógicas entre conceptos), después de la que se pedía a los sujetos que hiciesen un relato retrospectivo de sus procesos cognitivos durante la realización de la tarea (op. cit. p.97). Tareas semejantes se repetían muchas veces para que los relatos retrospectivos pudieran corregirse, corroborarse o ampliarse. Esos relatos eran frecuentemente suplementados por preguntas que dirigían la atención del sujeto a puntos particulares (ibid.).

Wundt también usaba la introspección en sus investigaciones, sin embargo, raramente del tipo cualitativo en el que el sujeto describe sus experiencias mentales. La especie de relato introspectivo que Wundt buscaba en su laboratorio trataba de modo especial de los juicios conscientes del sujeto acerca del tamaño, de la intensidad y de la duración de varios estímulos físicos -- juicios cuantitativos típicos de la investigación psicofísica -- (op. cit. p. 83).

Él creía en el estudio de la experiencia consciente tal como ocurría, no en la memoria de la misma después de haber ocurrido. Para él, los psicólogos deberían ocuparse del estudio de la experiencia inmediata, no de la mediata. Esto se debe a que consideraba la experimentación científica imposible en el caso del estudio de los procesos mentales superiores como el aprendizaje, el pensamiento, la memoria : estos procesos, por estar condicionados por hábitos lingüísticos y otros aspectos culturales, sólo podían estudiarse efectivamente mediante los enfoques no experimentales de la sociología, de la antropología y de la psicología social (op. cit. p. 81).

La investigación de Wundt se restringía, entonces, a los procesos mentales más simples, como la sensación y la percepción. Külpe procuró perfilar esa limitación.

Según Schultz y Schultz (1995, p. 98) :

“Külpe no rechazó el foco de Wundt sobre la experiencia consciente, el instrumento de investigación que era la introspección ni la tarea fundamental de analizar la conciencia en sus elementos. El blanco de su trabajo era expandir la concepción de objeto de estudio de la

psicología de Wundt con el fin de incluir los procesos mentales superiores, así como perfeccionar el método de introspección”.

Wundt defendía el punto de vista de que la experiencia consciente podía reducirse a sus elementos sensoriales o imagísticos componentes. Para él, toda la experiencia consciente se componía de sensaciones o imágenes. Külpe, a través de su introspección experimental sistemática, encontró evidencias en la dirección opuesta, i.e., el pensamiento puede producirse sin contenidos sensoriales o imagísticos (pensamiento sin imágenes) (op. cit. p. 98).

Karl Bühler, también de Würzburg, usó en sus investigaciones un método que era prácticamente el mismo que el de Külpe y que también recuerda mucho a la metodología usada hoy en los estudios sobre modelos mentales en el aprendizaje de las ciencias (ibid.) :

*“Su método de investigación suponía la presentación al sujeto de una cuestión que exigía cierta reflexión antes de poder responderse. Se pedía a los sujetos que hicieran el relato más completo posible de las etapas involucradas en la formulación de la respuesta, mientras el experimentador intercalaba preguntas sobre el proceso”.*

Los resultados obtenidos por Bühler reforzaron los descubrimientos de Külpe sobre los aspectos no sensoriales de la conciencia. (La existencia de tales procesos vino a ser, posteriormente, un presupuesto básico de la Psicología Cognitiva)

Por lo tanto, la metodología de la investigación en representaciones mentales, más específicamente en modelos mentales, no es nueva. Viene usándose en la psicología experimental desde el final del siglo pasado. Pero, como advierten Simon y Kaplan (1989, p. 21) :

*“A pesar de que los relatos verbales remontan a los introspeccionistas, el uso de tales relatos como datos no debe confundirse con la introspección. La introspección tomaba las verbalizaciones de los sujetos con valor terminal, constituyendo una teoría válida sobre sus propios procesos de pensamiento. Hoy, sin embargo, el análisis de protocolos trata los relatos verbales como fuente de datos de los que una teoría generada por el investigador debe dar cuenta -- tal teoría tal vez tenga la forma de una simulación computacional”.*

## **La cuestión de la conciencia y la computabilidad**

La sección anterior nos remite inevitablemente a la cuestión de la conciencia que, dentro del referente de la Psicología Cognitiva, nos conduce al tema de la computabilidad.

¿Al final, los modelos mentales son conscientes? ¿No conscientes? ¿Parcialmente conscientes?

¿Por qué esa insistencia con la simulación computacional, “rodar” el modelo, procedimiento efectivo, en fin, la metáfora del ordenador?

Empecemos por la segunda cuestión cuya respuesta parece ser más inmediata.

Modelos mentales es un tópico de la Psicología Cognitiva, que no debe confundirse con otras áreas de la Psicología o con otras psicologías. La Psicología Cognitiva es reciente, data de los años cincuenta de este siglo. En el núcleo duro de esta “nueva psicología” están *la existencia de*

*procesos mentales de “arriba hacia abajo” y la metáfora del ordenador.* Hacer Psicología Cognitiva implica aceptar que existen otros procesos mentales, además de aquellos “de abajo hacia arriba” que caracterizan el procesamiento de la información recibida a través de los órganos de los sentidos. (Ver referencia a las investigaciones de Wundt y Külpe en la sección anterior). Implica también pensar la mente como un sistema de cómputo, lo que no significa que sea un ordenador tal como lo conocemos, ni que las personas sean computables. Probablemente, hay aspectos de la mentalidad humana que nunca podrán ser explicados por una teoría científica. Sin embargo, hay otros que sí. Posiblemente, hay aspectos de la mentalidad humana que nunca podrán implementarse en un programa de ordenador. Pero hay otros que sí. Y ahí entramos en la Psicología Cognitiva, que pretende estudiar la mente humana científicamente.

Para Johnson-Laird (1983, p. 8), cualquier teoría científica de la mente debe, necesariamente, tratarla como un sistema de cómputo ; debe restringirse a aspectos que puedan formularse como programas de ordenador. “Abandonar este criterio es permitir que las teorías científicas sean vagas, confusas y, así como las doctrinas místicas, propiedad sólo de sus proponentes” (Ibid.). (El problema con este criterio es que los ordenadores actuales son seriales mientras que la mente es un sistema de procesamiento en paralelo).

Como se ha dicho, la metodología usada en la investigación sobre modelos mentales trata los protocolos verbales y otros documentos simbólicos o pictóricos producidos por los sujetos como *fuentes de datos* que deben ser explicados por teorías generadas por el investigador. Obviamente, para que tales teorías no sean vagas, confusas, místicas, deben poder describirse en forma de *procedimientos efectivos*. Procedimiento efectivo es aquel que puede ser ejecutado por una máquina (un ordenador) sin que ninguna decisión se tome basándose en la intuición o en cualquier otro ingrediente “mágico”.

Por eso, en la óptica de Johnson-Laird, el criterio de validez de los resultados de la investigación sobre modelos mentales es la posibilidad de describirlos en forma de procedimientos efectivos que puedan implementarse en ordenadores.

Los modelos mentales están en las cabezas de las personas. Si, a través de la investigación, conseguimos identificar algunos de estos modelos es porque interpretamos los datos (que están en los protocolos) a la luz de alguna teoría que tenemos. Además, el propio modelo identificado es, por sí solo, una teoría que tenemos sobre lo que está en la mente del otro. La manera de verificar la validez de esa teoría es describirla en la forma de un procedimiento efectivo que pueda formularse con un programa de ordenador. Naturalmente, la “rodaje” de ese programa debe proporcionar resultados compatibles con las previsiones de la teoría.

Pasemos ahora a la cuestión de la conciencia que, en la visión de Johnson-Laird tiene que ver con la computabilidad.

Johnson-Laird asume una posición funcionalista en relación con el problema cerebromente : *los fenómenos mentales no dependen de cómo está constituido el cerebro sino de cómo está funcionalmente organizado*, o sea, los fenómenos mentales corresponden a computaciones del cerebro (op. cit. p. 448 y 474). La metáfora del ordenador viene muy bien para la doctrina funcionalista : la mente es al cerebro como el “software” es al “hardware” en un ordenador. Pero hay un problema, tal vez el mayor de todos en lo que se refiere a fenómenos mentales : *la conciencia*.

Hasta hoy no se sabe realmente qué es la conciencia, qué hace, cuál es su función. Las alternativas a este problema van desde considerar a la conciencia como un fenómeno sobrenatural hasta suponer que es computable.

La alternativa funcionalista adoptada por Johnson-Laird es que *la conciencia depende de cálculos del sistema nervioso* (p. 450). Tales cálculos requieren un cerebro de cierto tamaño y complejidad porque la capacidad computacional depende de la memoria y porque la velocidad computacional depende del tamaño de las unidades que pueden procesarse simultáneamente, i.e., del número de procesadores que pueden ponerse en acción en una tarea (ibid.). Tanto la capacidad como la velocidad computacionales son importantes para la conciencia. Basta recordar el uso del lenguaje (que requiere gran habilidad mental computacional) y el hecho de que las decisiones conscientes se toman en tiempo real.

Otro punto importante en la visión funcionalista es *que el procesamiento mental ocurre en paralelo*. Tres argumentos sustentan esta afirmación (p. 451) :

1.- *La mente emplea distintos niveles de organización*; en el lenguaje, por ejemplo, en un primer nivel se procesan los sonidos del habla que, combinados, forman morfemas que tienen significados y deben ser procesados en otro nivel ; los morfemas combinados generan sentencias que implican otro nivel de procesamiento pues los significados de las sentencias van más allá de los significados de los morfemas ; por último, las inferencias a partir de los significados de las sentencias llevan a modelos mentales del discurso;

2.- *El procesamiento mental, en cada nivel, lleva implícito el contexto en consideración* ; de nuevo el lenguaje surge como ejemplo, pues la identificación e interpretación de las palabras depende del contexto ;

3.- *El procesamiento en diferentes niveles no es autónomo, pero sí interactivo* ; por ejemplo, los significados de las proposiciones pueden obtenerse por referencia a otras proposiciones ; el reconocimiento de una palabra facilita el reconocimiento de otras semánticamente relacionadas.

En resumen, la hipótesis es que hay diferentes niveles de organización mental porque los procesadores separados pueden operar en distintos niveles simultáneamente ; en un nivel concreto, un procesador trabaja cierto item mientras otros se encargan del contexto ; la comunicación entre procesadores en diferentes niveles permite que interactúen. Todo esto conduce a un aspecto esencial del procesamiento mental : ocurre en paralelo (p. 452).

Naturalmente, hay distintas maneras de implementar un procesamiento en paralelo. Una posibilidad es un sistema completamente distribuido en el que cada procesador está en pie de igualdad con los demás y no puede comunicar instrucciones a otros, solamente el producto del procesamiento realizado. Otra alternativa sería un sistema basado en un procesador central que “rodaría” un programa principal que daría instrucciones, y recibiría informaciones, de procesadores independientes.

Johnson-Laird imagina la mente humana funcionando en un sistema híbrido que saca partido tanto de la organización centralizada como de la distribuida : un procesador de alto nivel que monitoriza y controla los objetivos generales de los procesadores de nivel más bajo que, a su vez, monitorizan y controlan procesadores que operan más abajo que ellos, y así sucesivamente en una jerarquía de procesadores paralelos que, en su nivel más bajo, gobiernan las interacciones sensoriales y motoras con el mundo externo (p. 463).

El procesador de alto nivel correspondería a lo que en un ordenador se llama sistema operativo, un conjunto de programas que permite al operador humano manejar el ordenador. Cuando el ordenador está conectado, el sistema operativo se carga, automáticamente o a través de algún comando simple, y el operador tiene a su disposición mecanismos para recuperar programas y

archivos grabados en el disco duro, rodar los programas, leer los archivos, editar, grabar, imprimir, etc.

La mente tendría, entonces, un sistema operativo con considerable autonomía, pero sensible a demandas de otros procesadores y que podría ser conectado y desconectado por los mecanismos de control del sueño. Dependería de procesadores de segundo orden para percibir, entender, actuar, recordar, comunicar y pensar. Estos procesadores, a su vez, dependerían de procesadores de tercer orden para pasar “hacia abajo” instrucciones más detalladas de control y para pasar “hacia arriba” informaciones sensoriales parcialmente interpretadas. Habría interacciones entre procesadores del mismo nivel o de distintos niveles y mecanismos que permitiesen que los mensajes prioritarios de un nivel inferior interrumpiesen el procesamiento en trámite de un nivel superior (p. 464).

Una de las más importantes funciones del sistema operativo mental sería el desarrollo de nuevos programas para dar cuenta de nuevas situaciones, teniendo en cuenta que la mente humana puede desarrollar, rodar y modificar sus propios programas (modelos mentales).

Pero los “programas” que producen las personas para resolver problemas pueden tener grandes defectos (fallos, “bugs”), de modo que su implementación podría requerir configuraciones anómalas de los procesadores. Como el sistema operativo mental no tiene protección contra configuraciones anómalas, una forma primitiva de conciencia podría haber emergido originalmente de la red de procesadores paralelos como forma de perfilar tales configuraciones y otras interacciones patológicas entre ellos (ibid.).

Por lo tanto, en la base de las consideraciones puramente computacionales, Johnson-Laird argumenta que hay una división en la mente entre un sistema operativo de alto nivel y una organización jerárquica de procesadores. Y va más allá: él asume que “*los contenidos de la conciencia son los valores corrientes de los parámetros que gobiernan las computaciones de alto nivel del sistema operativo*” (p. 465). El sistema operativo puede recibir estos valores de otros procesadores, pero no puede inspeccionar las operaciones internas de esos procesadores. La selección natural habría asegurado que son necesariamente no-conscientes (p. 465).

El principal argumento de Johnson-Laird en este caso es exactamente el hecho de que hay cosas de la que el ser humano puede ser consciente y hay otras de las que no.

Cuando hablamos con otra persona podemos tener conciencia de las palabras que usa y si entendemos o no lo que está diciendo, pero no podemos volver conscientes los mecanismos que nos permiten entender las palabras y los significados de las proposiciones formadas; esto se procesa en un nivel no-consciente.

Podemos ser conscientes de muchas cosas -- sentimientos, actitudes, intenciones, motivos, expectativas, temores -- pero no de los mecanismos específicos subyacentes a las mismas.

Incluso en lo que se refiere a habilidades mentales, nunca seremos completamente conscientes de cómo ejercitarlas.

De manera análoga, *las personas no son conscientes de la naturaleza y de los mecanismos subyacentes a las representaciones mentales. Son conscientes de lo que está representado y de si es percibido o imaginado, pero no de la naturaleza intrínseca de la representación en sí* (ibid.).

En términos de modelos mentales, podríamos decir, entonces, que las personas son conscientes del contenido de sus modelos y de si son fruto de percepción o imaginación, pero no de sus mecanismos intrínsecos.

Es también importante advertir que, según la hipótesis de Johnson-Laird, cualquier intento de usar la introspección para hacer consciente alguna cosa que es no-consciente, fallará: no sólo porque la información es inaccesible sino también porque un proceso esencialmente paralelo tendría que captarse por las operaciones en serie del sistema operativo.

Hay una distinción entre las computaciones del sistema operativo que son en serie y las computaciones en paralelo de los múltiples procesadores. La introspección fuerza nociones intrínsecamente paralelas en un corredor serial.

Paradójicamente, la mente humana consciente, que es un sistema operativo en serie, no es capaz del grado de procesamiento paralelo necesario para producir conciencia. (p. 475).

Es también importante notar que además de no tener acceso a las operaciones internas de los procesadores de niveles inferiores, el sistema operativo mental tampoco tiene completo control sobre las mismas. Hay muchos fenómenos que sugieren que ciertos procesadores retienen bastante autonomía. Por ejemplo, los sentimientos de amor y odio pueden experimentarse conscientemente, pero no pueden ser invocados por una decisión consciente.

La mente intentando entender a la mente, i.e., buscando construir un modelo mental de un dispositivo que construye modelos mentales, es un problema que, además de ser paradójico, implica un enigma que es la conciencia.

Para Johnson-Laird, la construcción de un modelo mental es un proceso computacional e imagina que un componente de la solución de la paradoja cognitiva (i.e., la mente intentando entender a la mente) está en la construcción recursiva de modelos mentales, i.e., modelos dentro de modelos: “en un estadio 0, la mente construye un modelo de una proposición p; en el estadio 1, puede construir un modelo de sí misma operando en el estadio 0 y, en general, en cualquier estadio puede construir un modelo de sí misma operando en el estadio anterior” (p. 472).

Este procedimiento recursivo parece estar siempre presente en los fenómenos de la conciencia. Por ejemplo, un individuo puede percibir alguna cosa, tener conciencia de que está percibiéndola, tener conciencia de que es consciente de que está percibiendo esa cosa y así sucesivamente ... De manera análoga, la intencionalidad depende de la construcción recursiva de modelos dentro de modelos (“embedding of mental models”): un organismo puede tener una intención, i.e., una decisión consciente de actuar para alcanzar determinado fin, solamente si tuviera un sistema operativo capaz de construir un modelo mental de un estado de cosas futuro y, en un segundo estadio, capaz de construir un modelo de sí mismo construyendo el modelo mental inicial y decidiendo actuar de acuerdo con tal modelo (p. 473). Lo que se quiere decir es que una decisión consciente implica la construcción recursiva de modelos mentales. Un sistema operativo necesita, entonces, sólo tener acceso a un modelo de sí mismo con el fin de tener intenciones (ibid.).

Obviamente, las personas tienen intenciones y, por lo tanto, deben tener, y tienen, modelos mentales de sí mismas. Sus modelos incluyen recuerdos de cosas que sintieron o hicieron en el pasado y conocimientos de sus gustos y preferencias, habilidades y capacidades. No obstante, como ya se ha dicho, no tienen acceso al funcionamiento interno de los múltiples procesadores paralelos o del proceso que subyace a su propia representación. Sus modelos mentales de sí mismas están limitados a las opciones disponibles de sus sistemas operativos.

*Pero los modelos mentales no precisan ser ni completos ni enteramente perfeccionados para ser útiles.* En el caso de las personas, lo que su limitado conocimiento de su propio sistema operativo les da es un sentido de identidad, continuidad e individualidad (p. 474).

Para Johnson-Laird, la mente opera, entonces, en tres niveles computacionales principales : procesadores que computan en nivel no-consciente, mecanismos que construyen modelos mentales del mundo externo y dispositivos con la habilidad recursiva de construir modelos dentro de modelos.

El primer nivel es, por lo tanto, no-consciente, ¿Y los otros dos ? Bien, en el segundo, la hipótesis es que son conscientes el contenido de los modelos mentales y si son percibidos o imaginados, pero no lo son sus mecanismos intrínsecos. En el tercero, como recientemente se ha visto, la intencionalidad depende de la capacidad recursiva del organismo de construir modelos de sí mismo operando en estadios anteriores y la plena conciencia humana depende de la intencionalidad y de tener conciencia de estar consciente.

Esta respuesta no es totalmente esclarecedora. Y no podría serlo pues la conciencia es el mayor problema con el que se enfrenta la doctrina funcionalista, según la cual, los fenómenos mentales corresponden a computaciones del cerebro. Johnson-Laird intenta abordar este problema proponiendo la teoría que se ha presentado en esta sección, la cual supone que la mente está constituida por un sistema operativo de alto nivel y una organización jerárquica de procesadores. No obstante, como dice el propio Johnson-Laird (op. cit. p. 470), *la mente debe ser más complicada que cualquier teoría sobre la misma*.

### **Modelos mentales y memoria**

La teoría de los modelos mentales ha sido aplicada principalmente en el procesamiento del lenguaje, en la percepción de alto nivel y en el razonamiento. En el caso del lenguaje el foco primario de atención ha estado en cómo son construidos los modelos mentales cuando las personas comprenden lo que leen o lo que les es dicho (Garnham, 1997). Respecto a la percepción de alto nivel el interés de los investigadores se ha ocupado de los modelos mentales derivados de cómo las personas perciben el mundo, en particular los sistemas físicos, o sea, de representaciones internas de sistemas externos. En lo que se refiere al razonamiento se ha defendido la posición de que las personas utilizan modelos mentales para razonar y resolver problemas, no la lógica formal.

Considerando que todos esos aspectos de la cognición (lenguaje, percepción y razonamiento) dependen de algún tipo de memoria, la teoría de los modelos mentales debe, entonces, involucrar algunas suposiciones respecto a memoria. Según Garnham (op. cit. p. 169) los modelos mentales son modelos de situaciones en el mundo real o imaginario y mucho de lo que almacenamos en nuestra memoria de largo plazo es información sobre situaciones. La memoria de largo plazo también contiene información sobre individuos y cosas. Este tipo de información es utilizado en la construcción de modelos mentales. Por ejemplo, la comprensión de un texto implica la construcción de un modelo mental de su contenido. Sin embargo, la comprensión es un proceso constructivo que integra información sacada del texto e información almacenada en la memoria de largo plazo.

El proceso de construcción de modelos, y de integrar en ellos informaciones de la memoria de largo plazo, ocurre en un espacio mental que puede ser convenientemente llamado memoria de trabajo. Sin embargo, la teoría de los modelos mentales no está comprometida con ningún modelo particular de memoria de trabajo (ibid.).

Distintos énfasis en las memorias de trabajo y de largo plazo han sido utilizados por Keane, Byrne y Gentner (1997, apud Markman, 1999) para caracterizar la diferencia entre modelos mentales de razonamiento y modelos mentales de sistemas físicos. En el primer caso, los modelos

son contruidos para resolver problemas nuevos, pero dependiendo del problema la solución puede resultar muy difícil, o imposible, por requerir modelos múltiples al mismo tiempo en una limitada memoria de trabajo. En el segundo, el constructor del modelo debe tener un determinado conocimiento previo para que pueda de hecho modelar mentalmente el sistema físico. Es decir, este tipo de modelo enfatiza la manera cómo las personas almacenan y usan modelos de sistemas físicos sobre los cuales ya conocen alguna cosa. Este conocimiento previo está en la memoria de largo plazo.

En resumen, la teoría de los modelos mentales hace suposiciones sobre las memorias de largo plazo y de trabajo, pero no sobre cómo son esas memorias, cómo están estructuradas u organizadas. Las dos son simplemente instrumentales.

## Conclusión

En la investigación en la enseñanza de las ciencias, la década de los setenta fue la de las concepciones alternativas y la de los ochenta, la del cambio conceptual. Es muy posible que estemos hoy, en los años noventa, viviendo la década de las representaciones mentales, en particular, de los modelos mentales.

Los dos libros básicos sobre modelos mentales -- el de Johnson-Laird y el organizado por Gentner y Stevens -- son de 1983, pero fue en los primeros años de la década de los noventa, como bien destacan Pintó y col. (1996), cuando ese referente empezó a ganar cuerpo en el dominio conceptual de la investigación en la enseñanza de las ciencias.

No se trata, sin embargo, simplemente de una cuestión de moda, sino de una evolución, una etapa que es consecuencia de las anteriores. Con las investigaciones sobre concepciones alternativas, típicas de los años setenta, pasamos a saber, o pasamos a prestar atención a, que los alumnos vienen a la clase con una "ciencia alternativa". Luego, en seguida, pasamos a preocuparnos con cómo promover el cambio conceptual, i.e., cómo hacer que los alumnos abandonasen sus concepciones alternativas en favor de las concepciones científicas que, al final, eran mucho más explicativas y predictivas. Pero ahí nos perdimos porque supusimos que el cambio conceptual era una sustitución de una concepción por otra en la cabeza del alumno y que eso podría hacerse hasta con cierta facilidad, a través de estrategias instruccionales adecuadas. (Tales estrategias se basaban, generalmente, en el modelo kuhniano de cambio de paradigmas y en la idea piagetiana de conflicto cognitivo). Mucha investigación se hizo teniendo implícita esa suposición ; no obstante, la modestia de los resultados nos aclaró que el cambio conceptual es una cosa muy complicada y debe pensarse de otra manera, mucho más como una evolución conceptual que como una sustitución de concepciones (Moreira, 1994).

Vosniadou (1994), por ejemplo, interpreta el cambio conceptual como una modificación progresiva de los modelos mentales que el alumno tiene sobre el mundo físico, conseguida por medio de enriquecimiento o revisión. Enriquecimiento implica la adición de informaciones a los modelos existentes ; revisión supone cambios en las creencias o presupuestos individuales o en la estructura relacional del modelo.

De lo anterior se deriva que la aparición de un número cada vez mayor de artículos e investigaciones sobre modelos mentales se vea como una *consecuencia* del gran énfasis en el cambio conceptual que marcó mucho la investigación en enseñanza de las ciencias en la década pasada. Y tal vez ésa haya sido una etapa necesaria, pues fueron los resultados (como máximo modestos) de esa investigación los que llevaron a los investigadores a buscar otros referentes teóricos y, en esa búsqueda, llegar a los modelos mentales.

Posiblemente, éste es un referente teórico más prometedor, no obstante, más difícil desde el punto de vista metodológico. La idea de que las personas, o los alumnos en este caso, construyen modelos mentales del mundo, i.e., “re-presentan” internamente el mundo externo, es atrayente. El problema es que es difícil investigar tales modelos. Los modelos mentales de las personas, en vez de ser precisos, consistentes y completos, como los modelos científicos, son, simplemente, funcionales. En la investigación, en vez de buscar modelos mentales claros y elegantes, tenemos que procurar entender los modelos confusos, “contaminados”, incompletos, inestables que tienen realmente los alumnos. ¡Y eso es difícil !

En términos de teoría sobre modelos mentales, la de Johnson-Laird es, hasta hoy, la más completa y articulada. Por esta razón, impregnó todo este trabajo.

En lo que se refiere a la metodología, el análisis cualitativo de los protocolos verbales y de los documentos producidos por los alumnos ha sido la técnica más utilizada en la investigación sobre modelos mentales. A través de referencias históricas y descripción de investigaciones recientes, se intentó en este artículo prestar atención al aspecto metodológico del tema de los modelos mentales.

Además de lo expuesto, se procuró no dejar de lado la cuestión de la conciencia, una vez que es prácticamente imposible hablar de modelos mentales, o de cualquier teoría sobre la mente humana, si hablar de la conciencia. A propósito, para hablar de modelo mental es preciso tener un modelo mental de modelo mental, es preciso ser consciente de que se trata de un modelo mental de modelo mental, y por ahí vamos ... .

Todo esto, como se dijo al principio, se hizo con el objetivo de ayudar a la enseñanza y a la investigación en enseñanza de ciencias a la luz de ese (nuevo) referente que son los modelos mentales.

## **Bibliografía**

De Kleer, J. y Brown, J.S. (1983). Assumptions and ambiguities in mechanistic mental models. In Gentner, D. y Stevens, A.L. (Eds.). *Mental models*. Hillsdale, N.J. : Lawrence Erlbaum Associates, p. 155-190.

De Vega, M. (1984). *Introducción a la psicología cognitiva*. Madrid: Alianza Editorial. 562 p.

Eisenck, M.W. y Keane, M.T. (1994). *Psicología cognitiva : um manual introdutório*. Porto Alegre, R.S. : Artes Médicas. 490 p.

Garnham, A. (1997). Representing information in mental models. In: Conway, M.A. *Cognitive models of memory*. Cambridge, MA: The MIT Press.

Gentner, D. y Gentner, D.R. (1983). Flowing waters or teeming crowds : mental models of electricity. In Gentner, D. y Stevens, A.L. (Eds.). *Mental models*. Hillsdale, N.J. : Lawrence Erlbaum Associates. p. 99-127.

Greca, I.M. (1999). *Representaciones mentales*. Programa Internacional de Doctorado en Enseñanza de las Ciencias: Universidad de Burgos, España; UFRGS, Brasil. Texto de Apoyo n° 7.

- Greca, I.M. y Moreira, M.A. (1997). The kinds of mental representations -- models, propositions and images -- used by college physics students regarding the concept of field. *International Journal of Science Education*, London.
- Greca, I.M. y Moreira, M.A. (1996). Un estudio piloto sobre representaciones mentales, imágenes, proposiciones y modelos mentales respecto al concepto de campo electromagnético en alumnos de Física General, estudiantes de postgrado y físicos profesionales. *Investigações em Ensino de Ciências*, Porto Alegre, 1 (1) : 95-108.
- Gutiérrez, R. y Ogborn, J. (1992). A causal framework for analysing alternative conceptions. *International Journal of Science Education*, London, 14 (2) : 201-220.
- Halloun, I. (1996). Schematic modeling for meaningful learning of physics. *Journal of Research in Science Teaching*, 33 (9) : 1019-1041.
- Hampson, P. J. y Morris, P. E. (1996). *Understanding cognition*. Cambridge, MA : Blackwell Publishers Inc. 399 p.
- Harrison, A.G. y Treagust, D.F. (1996). Secondary students' mental models of atoms and molecules : implications for teaching chemistry. *Science Education*, 80(5) : 509-534.
- Johnson-Laird, P. (1983). *Mental models*. Cambridge, MA : Harvard University Press. 513 p.
- Johnson-Laird, P. (1996). Images, models, and propositional representations. In: De Vega et al. *Models of visiospatial cognition*. New York: Oxford University Press. p. 90-127.
- Lagreca, M.C.B. y Moreira, M.A. (1999). Tipos de representações mentais utilizadas por estudantes de Física Geral na área de mecânica clássica e possíveis modelos mentais nessa área. *Revista Brasileira de Ensino de Física*. No prelo.
- Mani, K. y Johnson -Laird, P. (1982). The mental representation of spatial descriptions. *Memory and Cognition*, 10(2) : 181-187. Apud Sternberg, R.J. (1996). *Cognitive psychology*. Forth Worth, TX : Harcourt Brace College Publishers. 555 p.
- Moreira, M.A. (1994). Cambio conceptual : crítica a modelos actuales y una propuesta a la luz de la teoría del aprendizaje significativo. Trabajo presentado en el "II Simposio sobre Investigación en Enseñanza de la Física", Buenos Aires, Argentina, 3 al 5 de agosto; en la Conferencia Internacional "Science and Mathematics Education for the 21 st Century", Concepción, Chile, 26 de septiembre a 1º de octubre; y en el Seminario Taller sobre Innovaciones en Enseñanza de las Ciencias Naturales en la Educación Básica, Montevideo, Uruguay, 24 al 28 de octubre.
- Moreira, M.A. y Greca, I.M. (1997). Concept mapping and mental models. *Meaningful Learning Forum*, 1 (1).
- Moreira, M.A. e Lagreca, M.C.B. (1998). Representações mentais dos alunos em Mecânica Clássica: três casos. *Investigações em Ensino de Ciências*, Porto Alegre, 3(2):
- Norman, D.A. (1983). Some observations on mental models. En Gentner y Stevens, A.L. (Eds.). *Mental models*. Hillsdale, NJ : Lawrence Erlbaum Associates. p. 6-14.
- Pintó, R. Aliberas, J. y Gómez, R. (1996). Tres enfoques de la investigación sobre concepciones alternativas. *Enseñanza de las Ciencias*, Barcelona, 14 (2) : 221-232.

Rodríguez Palmero, M.L. (1999). Identificación y tipificación de los modelos mentales relativos a la estructura y al funcionamiento celular construidos por estudiantes de Biología de COU. Proyecto de tesis doctoral en desarrollo.

Schultz, D.P. y Schultz, S.E. (1995). *História da psicologia moderna*. 7<sup>a</sup> ed. brasileira. São Paulo, Editora Cultrix. 439 p.

Simon, H.A. y Kaplan, A.C. (1989). Foundations of cognitive science. En Posner, M.I. (Ed.). *Foundations of cognitive science*. Cambridge, MA : The MIT Press. P. 1-47.

STAFF11 (1996). *Réprésentation de la connaissance*.  
<http://tecfa.unige.ch/staf/staf9597/beltrame/STAF11/concepts.html>

Sternberg, R.J. (1996). *Cognitive psychology*. Forth Worth, TX : Harcourt Brace College Publishers. 555 p.

Thagard, P. (1996). *Mind: introduction to cognitive science*. Cambridge, MA: A Bradford Book. The MIT Press.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, 4: 45-69.

Willians, M.D., Hollan, J.D. y Stevens, A.L. (1983). Human reasoning about a simple physical system. En Gentner, D. y Stevens, A.L. (Eds.). *Mental models*. Hillsdale, NJ : Lawrence Erlbaum.