



## **Modelos mentales, validez y credibilidad**

**Nilda J. Corral de Zurita**

### **Resumen**

Este artículo presenta resultados seleccionados de los obtenidos en un estudio cuyo objetivo general fue explorar algunos de los variados y complejos modos en que los contenidos y las formas se interpenetran en los procesos inferenciales. Se asume como marco teórico la teoría llamada de los *modelos mentales* o *teoría semántica del razonamiento* desarrollada por la psicología cognitiva. Se exponen las principales hipótesis acerca de la índole de los procesos inferenciales, sus operaciones, y la naturaleza de la competencia deductiva; al tiempo que se analizan las implicancias de este modelo teórico relativas a la racionalidad, a la vinculación de los procesos modelizadores con la lógica, y a la relación de la validez con la verdad. Los resultados propios que se seleccionan tratan los efectos de la credibilidad del contenido en el proceso inferencial mismo, en silogismos categóricos, y en relación a su influencia: a) en la inhibición de la búsqueda de modelos mentales alternativos; b) en la motivación para esta búsqueda. Estos resultados se comparan con los obtenidos en otros estudios en los que el efecto de las creencias no fue manipulado.

### **Introducción**

La psicología cognitiva desarrolla un amplio campo de estudios en el que se observa en situaciones controladas las respuestas de las personas en tareas de razonamiento. Sobre la base



de la evidencia empírica provista por estos estudios existe acuerdo entre los especialistas acerca de algunos rasgos fundamentales.

“En primer lugar la existencia de una competencia deductiva que se pone de manifiesto en la habilidad de algunos sujetos, en algunas ocasiones, para resolver problemas de razonamiento de singular dificultad, incluso con contenidos abstractos, alcanzando soluciones acordes con las previsiones de las teorías de la lógica. Ahora bien, esta competencia deductiva podríamos calificarla de “restringida” ya que la actuación de los sujetos en las tareas de razonamiento puede ser correctamente caracterizada por la gran cantidad de errores y sesgos que muestra. Por último, existe un patrón especialmente significativo en lo que se refiere a la influencia del contenido de la tarea o problema que se les presenta. En unos casos, la utilización de un contenido concreto y familiar puede conducir a la selección de conclusiones erróneas acordes con sus ideas previas, opiniones y actitudes; en otros, este tipo de contenidos puede facilitar la actuación normativa de los sujetos”\*.

Esto en cuanto al acuerdo que permite la evidencia empírica acumulada. Pero otra cuestión es cuál es la teoría que, utilizando un único conjunto de hipótesis, puede constituirse como un sistema de descripción, explicación y predicción de la competencia deductiva, los errores y sesgos, la influencia del contenido y el contexto en el razonamiento humano.

Al respecto, subsisten en el presente importantes desacuerdos entre los psicólogos cognitivistas. El núcleo de la polémica se centra en si lo que explica la competencia deductiva es la existencia de una capacidad básicamente lógica, que puede estar basada en reglas abstractas o en reglas dependientes del campo; o si razonamos mediante algún medio de naturaleza no lógica que, de algún modo racional, puede producir resultados correctos.

En la segunda posición se ubica el modelo teórico explicativo denominado *teoría de los modelos mentales* o *teoría semántica del razonamiento*, cuyo principal proponente es P. Johnson-Laird (1983, 1984, 1987, 1990, 1993). Como sus denominaciones lo indican se trata de una teoría del razonamiento humano que sostiene que no es necesario postular que nuestro sistema cognitivo contiene reglas de algún tipo para explicar el razonamiento, sino que afirma que tanto la competencia inferencial como los errores en la actuación y la influencia del contenido se explican de modo más plausible si se postula que el pensamiento opera a partir del significado de las situaciones descritas en los problemas.

La hipótesis de partida es que el razonador ordinario, aquel que no ha recibido formación sistemática en ninguna rama de la lógica, dispone de procedimientos modelizadores

\* García Madruga, J. A. (1995). Prólogo. En: Santamaría, C. Introducción al razonamiento humano.



de orden semántico para arribar a conclusiones correctas desde el punto de vista lógico (Johnson-Laird y Byrne, 1991). Adopta, por lo tanto, una perspectiva semántica de la lógica, sosteniendo que el razonamiento consiste en:

- a) operaciones semánticas sobre representaciones análogas en cuanto a su estructura a la situación descrita; esto es, que la estructura del modelo se corresponde en forma isomórfica con la situación descrita; y
- b) en un procedimiento de comprobación semántica del argumento; procedimiento consistente en que las personas hacen deducciones mediante la construcción de modelos mentales que representan analógicamente posibles estados del mundo y buscan contraejemplos mediante la construcción de modelos alternativos para asegurar la validez de sus conclusiones.

El individuo posee y aplica un principio semántico de validez que es suficiente para explicar la actividad inferencial: una inferencia es válida si la conclusión es verdadera en todas las situaciones en que las premisas lo son. En términos de proceso psicológico, una inferencia es válida si su conclusión no puede ser falsada por ningún modelo de las premisas.

Para la descripción del proceso inferencial el modelo teórico supone operaciones en secuencia que expondré con algún detalle oportunamente.

La teoría fue inicialmente propuesta para explicar la actuación en tareas con silogismos categoriales, posteriormente ampliada a todo el razonamiento deductivo y actualmente propuesta como una teoría general del razonamiento; incluido el razonamiento cotidiano. Es actualmente el modelo cognitivo que cuenta con mayor aceptación, tanto por las ventajas comparativas que presenta en lo referido a la explicación de la naturaleza del razonamiento y sus procesos como por el aceptable ajuste de sus hipótesis con resultados experimentales documentados. Pero no es una propuesta completa y uno de los aspectos que ha sido reiteradamente señalado es que a pesar de tratarse de una teoría semántica del razonamiento incorpora de modo limitado y parcial la complejidad de la influencia del contenido en la construcción y evaluación de modelos mentales.

### **Pensamiento y razonamiento**

Johnson-Laird (1990) ha propuesto una taxonomía del pensamiento cuyo punto de partida es la existencia o inexistencia de metas. Si el flujo de las ideas no tiene una meta es ensueño. Si tiene meta, el pensamiento puede ser determinístico o no determinístico. En el primer caso es cálculo, en el que una vez elegido un procedimiento ya no se tiene libertad sobre que hacer para obtener la respuesta correcta, porque cada paso está determinado por el estado



anterior del cálculo. Ensueño y cálculo marcan los extremos entre los que se ubica el pensamiento que tiene meta pero que no es determinístico. Si la meta no está definida con precisión estamos en el dominio del pensamiento creativo, en el que no hay sólo una respuesta correcta y no se sigue un procedimiento estrictamente determinado; tal como ocurre en una obra de arte o en la generación de una hipótesis científica, o en cualquier situación, aún trivial, en la que se intenta crear una idea nueva.

El razonamiento, en cambio, ocurre cuando el pensamiento tiene una meta precisa pero no cuenta con un procedimiento rutinario para llegar a ella. El razonamiento pone en relación unas premisas con una conclusión, pero la cualidad de esta relación da lugar a las dos variantes tradicionales: deducción e inducción. Para establecer la distinción entre estas formas de razonamiento el autor recupera el concepto de *información semántica*, definida del siguiente modo "cuanto más estado de cosas de entre los que pueden considerarse elimine una proposición, mayor información semántica contiene"(1990: 206). Por lo tanto, si una conclusión descarta más estado de cosas adicionales a los descartados por las premisas es inducción. El razonamiento inductivo es definido como un modo sistemático de razonar que aumenta en la conclusión la información contenida en las premisas.

Por oposición, la deducción es precisamente aquel razonamiento cuya conclusión no contiene más información semántica que las premisas. Es una inferencia válida ya que su conclusión es verdadera en cualquier situación en la que las premisas son verdaderas. Desde el punto de vista estrictamente lógico, solo la deducción dará lugar a conclusiones necesariamente válidas. Hasta aquí la taxonomía que permite caracterizar al razonamiento deductivo como una de las formas del pensamiento

#### *El silogismo categórico en los estudios sobre el razonamiento*

El objetivo de las investigaciones sobre el razonamiento es comprender la naturaleza de la actividad mental de los sujetos y el medio para estudiarla es la utilización de tareas diseñadas para tal propósito. Los estudios sobre razonamiento deductivo han utilizado profusamente el silogismo categorial debido a las ventajas que presenta. Se trata de tareas bien definidas y estructuradas que contienen la información necesaria y suficiente para elaborar la conclusión, facilitando en consecuencia la evaluación de los procesos racionales. Por tratarse de un razonamiento que trabaja con la estructura interna de los enunciados - los términos y relaciones que los componen- permite la modelización por componentes, facilitando seguir su modelización mental.

El silogismo categórico es una forma altamente circunscrita del razonamiento deductivo en el que las proposiciones son afirmaciones acerca de clases y las inferencias dependen de las



relaciones de inclusión o de exclusión entre ellas. Varían notablemente en la dificultad que presentan y dan lugar a gran variedad de respuesta.

Se define como una forma de razonamiento consistente en tres proposiciones categóricas, dos son las premisas y una es la conclusión, que contiene tres, y sólo tres, términos. La primera premisa o premisa mayor contiene el predicado de la conclusión, y la segunda premisa o menor contiene el sujeto de la conclusión, de acuerdo a lo prescrito por la lógica tradicional. El término que se repite en las dos premisas y no aparece en la conclusión es el *término medio* y su función es permitir establecer una relación entre el sujeto y el predicado que aparece en la conclusión. Es una inferencia mediata porque la conclusión fija la conexión entre dos términos sobre la base de sus relaciones con un tercero.

Tanto las premisas como la conclusión son enunciado que incluyen cuantificadores, que varían en cantidad (universal o particular) y en cualidad (afirmativa o negativa), dando lugar a cuatro tipos de enunciados a los que se acostumbra simbolizar con vocales mayúsculas:

A: Universal Afirmativa: Todo A es B - I: Particular Afirmativa: Algún A es B  
 E: Universal Negativa: Ningún A es B - O: Particular Negativa: Algún A no es B

El *modo* de un silogismo es determinado por la cualidad y cantidad de los enunciados categóricos utilizados en las premisas y la conclusión y el orden en que se presentan.

Los silogismos difieren también en la ubicación del término medio, determinando las cuatro *figuras* del silogismo. El término medio puede ser el término sujeto de la premisa mayor y el predicado de la premisa menor (primera figura), o puede ser el predicado en ambas premisas (segunda figura), o puede ser el sujeto de ambas (tercera figura), o puede ser el predicado de la mayor y el sujeto de la menor (cuarta figura).

Para representar la disposición de los términos en las cuatro figuras tradicionales del silogismo se acostumbra en psicología del razonamiento la utilización de las vocales mayúsculas: "A" para el término extremo de la primera premisa, "B" para el término medio, y "C" para el término extremo de la segunda premisa, considerados más neutrales que los tradicionales "S", "m", "P" que indican la función que cumplen ("S" y "P" sujeto y predicado en la conclusión; "m" término medio).

Fig. primera	Fig. segunda	Fig. tercera	Fig. cuarta
B - A	A - B	B - A	A - B
<u>C - B</u>	<u>C - B</u>	<u>B - C</u>	<u>B - C</u>
C - A	C - A	C - A	A - C



En el sistema lógico tradicional la conclusión en cada caso será C-A . Sin embargo, cuando la gente deriva una conclusión no toma necesariamente como sujeto de la conclusión el término extremo de la segunda premisa. Johnson-Laird y Steedman (1978; Johnson-Laird y Bara, 1984) eliminan la restricción de que el sujeto de la conclusión esté necesariamente en la premisa menor puesto que desde el punto de vista psicológico es irrelevante; decisión que tuvo una importante consecuencia. En efecto, al optar por el paradigma experimental de elaboración de la respuesta, en lugar del paradigma de evaluación de respuestas presentadas, y al eliminar la mencionada restricción, permitiendo a los sujetos obtener conclusiones en uno y otro sentido, se puso en evidencia la existencia del efecto de la figura y la fuerza de este efecto en las figuras primera y cuarta. Este efecto es de carácter estructural y consiste en que en la figura primera prevalecen decididamente las conclusiones C - A; en tanto en la cuarta la tendencia es hacia conclusiones A - C. Aunque está en discusión la naturaleza de este efecto, numerosos trabajos han corroborado su existencia y su importancia como predictor del orden en que la información será procesada. Este factor tiene efecto facilitador cuando el problema admite respuesta correcta en la dirección que marca la figura. Pero también, en el caso de la figura cuatro, el efecto de la figura supone una importante dificultad adicional cuando el modo del silogismo no admite respuesta correcta de tipo A - C. En las figuras segunda y tercera la figura no ejerce efecto por lo que no se predice una dirección definida en las conclusiones.

**Los estadios del proceso**

El proceso inferencial es descrito como una estrategia general organizada en estadios, aplicable a la habilidad inferencial general pero que aquí exponemos en su versión ajustada a la especificidad del razonamiento silogístico.

1- El razonador construye un modelo mental de la primera premisa. Estos modelos se constituyen de unidades simples que se componen para representar estados de cosas. En el caso de las expresiones cuantificadas del silogismo, el modelo contiene elementos que representan un número arbitrario, aunque pequeño por razones de economía cognitiva, de individuos representativos del conjunto y las ecuaciones que los ligan.

Así, de acuerdo a la última notación propuesta por Johnson Laird y Byrne (1991), las representaciones de los modelos mentales de enunciados en los cuatro modos del silogismo serán:

Todos los A son B	Algunos A son B	Ningún A es B	Algunos A no son B
[ a ]        b	a            b	[ a ]	a
[ a ]        b	a            b	[ a ]	a
....	....	[ b ]	a
		[ b ]	[ b ]
		.....	.....



El corchete indica que el conjunto está exhaustivamente representado, es decir que todos sus miembros lo son del otro conjunto; mientras que la ausencia del corchete indica que no lo está y que, por lo tanto, puede haber individuos que no sean miembros del otro conjunto. Las líneas de punto indican que podría haber otra clase de individuos que no se encuentran representados en el modelo. Es importante advertir que de este modo la representación admite la especificación de todas las maneras relevantemente diferentes en que las premisas de un silogismo pueden ser verdaderas.

Como ejemplo, para la afirmación "todos los A son B" el modelo muestra que no hay individuos "a" que no sean también individuos "b"; que puede haber, aunque no necesariamente, individuos "b" que no sean individuos "a". El modelo también conserva la posibilidad, indicada por la línea de puntos, de la existencia de individuos que no son "a" ni "b". Pero también es claro que prohíbe la interpretación de la existencia de individuos "a" que no sean "b".

Por ejemplo, para el enunciado Todos los gorriones son pájaros, el conjunto de los gorriones se encontrará exhaustivamente representado ya que cada uno de sus elementos pertenece al conjunto de los pájaros; en cambio el conjunto de los pájaros no estará exhaustivamente representado porque es posible que pueda haber pájaros que no sean miembros del conjunto gorriones, como efectivamente es el caso.

2- A continuación construye un modelo mental de la segunda premisa y añade esta información al modelo mental de la primera, combinando ambas interpretaciones en un único modelo y teniendo en cuenta los modos alternativos en que puede hacerse. Así, si el primer enunciado es " Todos los A son B" y el segundo "Todos los B son A", el modelo integrado resultante será:

{	[a]	b]	c
[	[a]	b]	c
.....			

Se asume que en la combinación de premisas el enlace entre los ítems extremos se establece a través del término medio, tanto con premisas positivas como negativas.

En el ejemplo presentado la estructura del problema permite el enlace directo por la disposición de contigüidad del término medio. Cuando la estructura no lo permite se requiere de operaciones adicionales para hacer coincidir los términos medios. En consecuencia la figura de las premisas afecta la integración de la información.



Los autores sostienen que existe una preferencia de los sujetos a construir primero un modelo de la primera premisa para integrar en él la segunda. Pero cuando la figura del silogismo no permite esta operación, el razonador tendrá que producir cambios en sus representaciones iniciales. En los silogismos en la figura primera: B-A, C-B, necesitarán invertir el sitio de las premisas, iniciando el procesamiento de la información a partir de la segunda premisa. La figura segunda y tercera requieren invertir espacialmente el modelo de alguna de las premisas, es decir cambiar la ubicación de los elementos correspondientes al sujeto y al predicado. Esta operación de dar la vuelta a la interpretación de una premisa no implica, en principio, conversión ilícita, puesto que se ve afectado sólo el orden de la accesibilidad en la memoria operativa, dando lugar a una nueva interpretación lógicamente correcta. Como ejemplo, si es necesario dar la vuelta a la interpretación de una premisa del tipo Todos los A son B, dará lugar al siguiente modelo:

Todos los A son B		Algunos B son A	
[a]	b	b	[a]
[a]	b	b	[a]

Por el principio según el cual el modelo se lee en el orden en que se organizaron las premisas, el orden de los términos en la conclusión que se construye dependerá de las operaciones que se describieron, permitiendo explicar y predecir la mencionada tendencia hacia conclusiones de tipo A-C en la figura cuarta y del tipo C-A en la figura primera.

Estas operaciones también permiten predecir que la figura es un factor que interviene en la mayor o menor dificultad del problema: a mayor número de manipulación del modelo mental necesarias para la contigüidad del término medio, mayor dificultad. La figura cuarta será la más fácil, seguida de la figura primera. Las figuras segunda y tercera serán más difíciles.

3- Una vez combinadas las premisas en un modelo simple el sujeto formula una conclusión inicial. Este estadio se corresponde con un proceso de descripción verbal que expresa una relación que no está explícita en las premisas. Construye una conclusión inicial que verbaliza la relación, si la hubiere, entre los términos extremos.

4- A continuación inicia la búsqueda de modelos alternativos - o contraejemplos- que puedan refutar la conclusión que ha elaborado. Si ocurre que la información de las premisas puede ser representada por más de un modelo, el razonador opera combinando modelos que puedan expresar todas las situaciones posibles descritas en las premisas; es decir, vuelven a interpretar y combinar los significados de las premisas. En caso de no encontrar una conclusión alternativa considera válida la conclusión inicial. En caso de encontrarla, intentará comprobar si hay una conclusión verdadera para todos los modelos que haya podido construir.





no serlo en otras ( Johnson-Laird, 1984).

Asumir la tesis de la racionalidad variable permite tratar el pensamiento como habilidad que se adquiere mediante la experiencia y prácticas reales, y explicar las sensibles diferencias observadas entre sujetos relativas a la capacidad que muestran para resolver correctamente inferencias difíciles.

En efecto, algunos problemas inferenciales - los considerados de un modelo- no requieren habilidades específicas y son ejecutados correctamente por la mayoría de los sujetos. Pero otros de mayor complejidad sí las requieren- los considerados de más de un modelo- y en ellos se pone de manifiesto la existencia de notables diferencias en la actuación de los sujetos. Johnson-Laird ( 1984) sostiene que las diferencias individuales en la capacidad para hacer inferencias, al menos en aquellas situaciones en las que el sujeto dispone de la información necesaria, depende de tres habilidades:

a) la capacidad de construir modelos mentales de las situaciones descritas en las premisas; b) la conciencia del principio semántico de refutación para buscar diferentes modelos de las mismas premisas, acompañada de la habilidad para ponerlo en práctica en la búsqueda de diferentes modelos de las mismas premisas; c) la capacidad para verbalizar adecuadamente la característica común de un conjunto de modelos mentales.

En síntesis, de acuerdo a este modelo explicativo las personas no poseen una competencia deductiva subyacente basada en reglas y el sistema cognitivo humano para razonar opera sólo con principios y procedimientos semánticos. Aunque su puesta en práctica es muy variable, estos principios y procedimientos constituyen el común denominador de la racionalidad.

#### *Procesos modelizadores y lógica*

Si se aceptan estas hipótesis acerca de la naturaleza de la competencia deductiva y de la racionalidad, cabe la pregunta: qué relación guardan entonces el pensamiento y la lógica?. Es evidente que no se otorga a la lógica carácter descriptivo alguno en relación al pensamiento puesto que éste no opera con sus reglas y principios. Tampoco tendría carácter preceptivo, en el sentido de que la corrección del razonamiento se asegure por su adecuación a los principios lógicos ya que la mente no posee estos principios.

“La adquisición de la competencia deductiva es mucho menos problemática si lo que debe adquirirse es cierta capacidad para construir modelos del mundo, ya sea directamente mediante la percepción o indirectamente mediante la comprensión del lenguaje, y cierta capacidad para buscar modelos alternativos” (Johnson-Laird, 1991: 114)



Sin embargo, puede entenderse que la teoría no obliga a divorciar los procesos modelizadores de la lógica puesto que, si bien estos procesos son de naturaleza estrictamente semántica, puede alcanzarse el razonamiento correcto si los procedimientos semánticos de construcción y de comprobación de modelos mentales son adecuados y exhaustivos. Aunque el proceso es falible porque es susceptible de producir falsos resultados, y de hechos los produce, el sujeto dispone de un procedimiento semántico para asegurar la validez de sus razonamientos. Este procedimiento es, en definitiva, una prueba lógica a la que se somete la representación inicial.

Si la validez reside en las posibles interpretaciones del argumento y no en las propiedades y relaciones formales, parece entonces que la validez de un argumento, en tanto consistente en su carácter preservador de la verdad, ha sido aquí convertida en una propiedad psicológica. Si así es, las respuestas de los sujetos pueden ser valoradas según los estándares lógicos de la inferencia deductiva y ser calificadas como correctas o incorrectas, tal como típicamente lo hacen las aplicaciones experimentales realizadas en el marco de la teoría.

#### *Verdad y Validez: componentes de la adecuación deductiva*

El razonamiento es el proceso cognitivo que nos permite pasar de una información dada a otra. En el razonamiento deductivo se parte de unas premisas, información dada, para alcanzar una conclusión que se siga necesariamente de las mismas. La noción de necesidad lógica de la conclusión, o validez, es la característica distintiva de este razonamiento. Este carácter necesario se debe a que *no aporta nueva información* en la conclusión, en el sentido de proporcionar conocimiento fáctico nuevo. En cambio, reordena la información existente revelando una relación que no se encontraba explícita en las premisas. Es conocimiento sólo en el sentido de poner de manifiesto información que ya está contenida en otra; de modo tal que a partir del contenido total de la información provista por las premisas, extrae o explicita una parte del mismo. Es la razón por la que esta forma inferencial garantiza la validez de la conclusión.

Entiendo que esta caracterización del razonamiento deductivo es equivalente a las definiciones dadas por Johnson-Laird y colaboradores; la ya mencionada "razonamiento que no incrementa la información semántica en la conclusión", o bien "deducir es mantener información semántica, simplificarla y alcanzar una nueva conclusión" (Johnson-Laird y Byrne, 1991:22), definiciones que le permiten diferenciarlo del razonamiento inductivo. La validez es una propiedad de los razonamientos que depende de su forma o estructura y que asegura que no existe un estado posible de cosas en el que sienta verdaderas las premisas la conclusión sea falsa.

Es por esta propiedad que podemos razonar deductivamente a partir de proposiciones



que sabemos verdaderas, pero también a partir de conjeturas y aun de contenidos arbitrarios o de símbolos. Sólo en la primera situación la validez garantiza la verdad de la conclusión, no así en la segunda situación puesto que mientras no se determine el valor de verdad de las premisas no se determina tampoco el valor de verdad de la conclusión. La verdad es una propiedad de las proposiciones, no de los razonamientos. Por lo tanto, el carácter necesario de la verdad en la conclusión exige la validez del razonamiento y la verdad de las premisas.

"Que el argumento es válido significa que las premisas apoyan o justifican la conclusión, en el sentido de que *caso de estar las premisas justificadas*, la conclusión queda también justificada; esto es, los argumentos válidos " trasladan " la justificación de las premisas a la conclusión. Por tanto, aunque la conclusión se infiera efectivamente de las premisas, puede carecer de justificación si alguna de las premisas carece de ella. La validez de un argumento no justifica *por sí sola* la conclusión" (Diez y Moulines, 1997: 37).

Los enunciados asertóricos, cuya finalidad es describir como son las cosas, son verdaderos si las cosas son efectivamente como se afirma, y este es el tipo de enunciados que se presenta en las situaciones experimentales que utilizan contenidos temático; en consecuencia la derivación lógica ya no es solo cuestión de formas sino que en ella intervienen los conocimientos y las convicciones de las personas.

Para una teoría como la que estamos tratando, que abandona la aproximación al razonamiento deductivo mediante reglas de inferencias para reemplazarla por principios semánticos y operaciones sobre la semántica de las representaciones, las cuestiones vinculadas a los contenidos adquieren especial importancia y no pueden ser soslayadas.

#### *Creencias y validez*

Los estudios experimentales sobre el silogismo categórico se han centrado prioritariamente en el estudio de los efectos estructurales de la figura y la forma. En cuanto a la influencia del contenido temático en la actuación de los sujetos ha sido generalmente tratado: o bien comparativamente con contenidos abstractos para testear posibles efectos facilitadores; o bien como una consecuencia de la credibilidad de los enunciados. Respecto a este último tratamiento la teoría de los modelos mentales postula que la credibilidad puede afectar el razonamiento en los tres instancias del proceso ( Oakhill, Garnham y Johnson-Laird, 1990) .

En la instancia interpretativa puede distorsionar la interpretación de las premisas modificando su significado de acuerdo a los conocimientos previos. En la segunda instancia puede sesgar el proceso de deducción, de modo tal que cuando los sujetos obtienen una conclusión tentativa creíble se esforzarán menos en buscar contraejemplos; e inversamente, cuando obtienen una conclusión increíble puede influenciar motivando la búsqueda de modelos



y constituye también un espacio de juego donde pueden introducirse la intervención y el cambio y donde, en cada instante, se funda de nuevo el vínculo social” (ibid).

Ahora bien, en el grupo áulico, se reconocen dos tipos básicos de interacción:

- “aquella que es mediatizada por el docente en torno a la actividad de enseñanza-aprendizaje escolar (interacción didáctica docente-alumnos) y
- aquella que resulta de la dinámica propia del grupo de pares (estructura grupal) (Roselli, 1990, b: 5).

Estos tipos básicos de interacción tienden a corresponderse con la idea del grupo-clase y el grupo-curso, respectivamente.

En el estudio del primer tipo básico de interacción (las **interacciones docente – alumnos**), se han destacado distintos “procesos interactivos observables” en el grupo-clase. Por ejemplo, se han podido distinguir “los procesos de interdependencia de roles”, que describen cómo cierta conducta “dominadora” del maestro podría generar pasividad, ausencia de cooperación y agresividad en los alumnos ; y cómo una conducta “integradora” provocaría comportamientos amistosos, activos y cooperativos. Además, se han considerado “los procesos de dependencia” (funcional y/o afectiva), observable en cualquier grupo que presente una estructura asimétrica de poder (como sería el caso de una clase). De esta forma, se pudo comprender cómo se regulan las relaciones entre docentes y alumnos, además, las oposiciones, los bloqueos u obstáculos de aprendizaje entre otros. También se ha estudiado como el tipo de liderazgo ejercido por el docente (autocrático, democrático y “laissez faire”), puede condicionar la eficacia de la acción educativa. (Gilly, 1984)

En el estudio del segundo tipo básico de interacción (la **interacción entre pares**), se ha constatado que según como sean los tipos de interacciones entre los alumnos (ya sean cooperativas, egocéntricas o asimétricas) se producirán distintos rendimientos escolares en los mismos (Roselli, 1994). Un número importante de teorizaciones e investigaciones empíricas están demostrando los beneficios que conlleva el aprendizaje cooperativo en los ambientes educativos (Echeita, 1995). Específicamente, en el estudio de este tipo de interacción, se pueden encontrar una serie de trabajos conocidos por sus implicancias educativas, por un lado, tenemos la línea neopiagetiana del conflicto sociocognitivo (por ej., con los trabajos de Perret Clermont) y por otro, la línea Vygotskiana<sup>60</sup>. En ambas propuestas, se destaca la idea de que la interacción

---

el vínculo”, los cuales no implican necesariamente una presencia conjunta de los sujetos, ya que pueden mantenerse y proseguirse a distancia, como por ej. en una relación profesional (Ibid: 15).

<sup>60</sup> Estudios específicos en estas líneas de investigación pueden encontrarse en Fernández Berrocal y Melero Zabal, 1997 (ver Bibliografía).



alternativos; lo que permite predecir que el efecto de la credibilidad de la conclusión será mayor en los problemas de varios modelos, puesto que en los de un modelo no existen contraejemplos. En la instancia de selección de la conclusión, la credibilidad puede operar como filtro; de modo tal que las personas no tendrían en cuenta conclusiones poco creíbles substituyéndolas por otras creíbles, o bien concluyendo que no se deriva nada.

Al respecto, se ha reunido amplia evidencia experimental que prueba que las creencias y el conocimiento previo afectan a los sujetos cuando juzgan si cierta conclusión deriva de las premisas que se les presentan. En el apartado siguiente se presentan resultados propios relativos a los efectos del sistema de creencias en el proceso inferencial.

### **Situaciones inferenciales y resultados**

Las situaciones inferenciales y los resultados que se exponen a continuación forman parte de un estudio cuyo objetivo general fue explorar la penetrabilidad de los procesos de razonamiento a los factores de interpretación, para la posterior formulación de hipótesis relativas a los principios que regulan la interacción de formas y contenido. Corresponden a una muestra de 28 estudiantes de primer año de las distintas carreras de la Facultad de Humanidades; se utilizaron tareas de construcción de la conclusión; y la selección que se presenta está referida a:

1- la influencia de la credibilidad de los contenidos en el proceso inferencial mismo: a) inhibiendo la búsqueda de modelos mentales alternativos; b) motivando esta búsqueda.

#### ***1- Sesgo de las creencia***

Buena parte de la investigación sobre la influencia de los contenidos se ha centrado en el análisis de la interacción validez/ credibilidad para determinar como opera el denominado "sesgo de las creencias". El sesgo consiste en que las personas aceptan con más facilidad aquellas conclusiones que están de acuerdo con sus creencias que aquellas que no lo están, por lo que las respuestas están determinadas más por la credibilidad de la conclusión que por la validez lógica (Velazco y García Madruga: 1997). Por lo tanto, puede anticiparse que cuando se comparan problemas de conclusión creíble e increíble, la tendencia será hacia el procesamiento completo de todos los modelos posibles en los increíbles y no en los creíbles.

Entre los estudios que manipulan experimentalmente la credibilidad y la validez se encuentra el de Evans, Barston y Pollard (1983), en el que los sujetos tenían que juzgar la validez de una y solamente una conclusión. Los resultados que obtuvieron mostraban que los



silogismos válidos eran aceptados con más frecuencia que los inválidos; que los silogismos creíbles eran aceptados con más frecuencia que los increíbles; y, lo más importante, que el sesgo que introducen las creencias es más acusado para los silogismos inválidos que para los válidos.

**a) credibilidad de los enunciados y validez**

En los materiales de prueba incluí cuatro silogismos con contenido creíble y sin conclusión proposicional válida, utilizando premisas acordes con las creencias de los sujetos y de las que puede obtenerse conclusiones compatibles con ellas y altamente plausibles. Por el carácter creíble de las premisas, puede suponerse que el sesgo de las creencias no actúa en la instancia interpretativa. En cambio, actuaría en el proceso inferencial mismo desalentando la búsqueda de modelos mentales que pudieran falsar la primera conclusión derivada.

Los resultados son indicativos de la tendencia a privilegiar la credibilidad de los contenidos para aceptar conclusiones que no son válidas pero que están en concordancia con las creencias, casi sin tener en cuenta la corrección o incorrección de las inferencias que intervienen. Expresan una conducta orientada a establecer relaciones que se consideran verdaderas más que a determinar relaciones lógicas entre enunciados.

Estos resultados se comparan con los obtenidos por Johnson-Laird y Bara (1984) en silogismos de la misma forma, en los que se utilizan contenidos familiares pero neutros en relación a los sistemas de creencias de los sujetos ( como ejemplo: enunciados que relacionan arqueólogos, biólogos y jugadores de ajedrez; en un contexto que los ubica en una habitación)

1- Un silogismo en la segunda figura y modo AA, de dos modelos mentales y sin conclusión válida, con las siguientes premisas:

*Todos los niños son curiosos/ Todas las personas inteligentes son curiosas*

El primero modelo que se construye induce la conclusión " todos los a son b", en nuestro caso " todos los niños son personas inteligentes"; pero también la conclusión " todos los b son a", que en nuestro caso es " todas las personas inteligentes son niños". Como en esta forma silogística la figura no ejerce influencia sobre la dirección de la respuesta los sujetos pueden obtenerla en una u otra dirección. Pero con las premisas utilizadas la segunda conclusión no es empíricamente verdadera y ningún sujeto la presentó.

[ niños]	curiosos	[personas inteligentes]
[ niños]	curiosos	[personas inteligentes]



Si el sujeto advierte que la extensión del término medio no está determinada en ninguna de las premisas puede construir un segundo modelo mental que corresponde a una nueva interpretación de las premisas:

[ niños]	curiosos		
[ niños]	curiosos		
	curiosos	[personas inteligentes]	
	curiosos	[personas inteligentes]	

Al evaluar los dos modelos advertirá que no hay una conclusión que sea compatible con ambos y concluirá que no se deriva nada.

Pero solamente un sujeto presentó esta conclusión. Los demás obtuvieron conclusión proposicional, predominando la respuesta "todos los niños son personas inteligentes" ( 60%). Otro tipo de respuesta rescató la relación causal entre los términos del problema: " los niños son inteligentes porque son curiosos".

En cambio, el mismo silogismos con contenidos familiares pero "neutros" obtuvo en el estudio que utilizamos para comparar 35% de respuestas " no hay conclusión válida", y 45% de respuestas del mismo tipo que la nuestra: " todos los A son B".

2- Un silogismo de la primera figura y modo IA, de dos modelos y sin conclusión valida, con las siguientes premisas:

*Algunos felinos son peligrosos /Todos los tigres son felinos*

Por tratarse de la primera figura, de acuerdo al proceso de construcción de modelos que presentamos al inicio de este trabajo, requiere rehacer la interpretación de las premisas para ubicar en primer lugar el modelo mental de la segunda e integrar el modelo de la primera premisa. El primer modelo mental integrado resultante será:

[tigres]	felinos	peligrosos
[tigres]	felinos	peligrosos

Al igual que en el caso anterior, si el sujeto advierte que el alcance del termino medio no está determinado puede construir un segundo modelo que hace falso al primero. Al evaluarlos advertirá que no hay conclusión compatible con ambos y concluirá que no se deriva nada.

[tigres] felinos



[tigres] felinos  
          felinos peligrosos  
          felinos peligrosos

Sólo un sujeto consignó que no se puede derivar nada y otros dos expresaron que advertían que no se deriva conclusión válida, pero que igual la escribían "por convicción". Los demás presentaron conclusión proposicional, predominando la respuesta "todos los tigres son peligrosos" ( 77%), que es la respuesta que induce la figura y que se corresponde con las creencias de los sujetos.

Johnson -Laird y Bara obtienen: 20% de respuestas "no hay conclusión válida" y 75 % de respuestas de tipo C-A. acordes con el efecto de la figura.

3- Un silogismo en la segunda figura y modo II, de dos modelos y sin conclusión válida, con las siguientes premisas.

*Algunos animales son agresivos/ Algunos monos son agresivos.*

Tres sujetos responden que no hay conclusión válida, en tanto un 55% consigna esta respuesta en un silogismo del mismo modo y figura en el estudio de Johnson-Laird y Bara. Interpreto que este alto porcentaje se debe a que el primer modelo mental integrado mantiene indeterminado el alcance de todos los términos e induce la búsqueda de un modelo alternativo, mientras que en nuestro caso la credibilidad de la conclusión que admite este modelo induce a no continuar la búsqueda.

Como no actúa el efecto de la figura los sujetos pueden sacar conclusión en una u otra dirección. Con los contenidos que estamos considerando 33,50 % de los sujetos obtuvieron la conclusión " algunos animales son monos".

En tanto otro 30% produjo la respuesta " todos los monos son animales". Ningún modelo mental que pueda ser construido a partir de estas premisas sostiene esta respuesta, por lo que puede entenderse que quienes la dieron optaron por organizar el modelo partiendo de la segunda premisa, obteniendo en consecuencia la conclusión " algunos monos son animales" que es empíricamente falsa. En este caso puede suponerse que el efecto de las creencias se ubica en el mecanismo de "filtrado" de una conclusión no concordante con el conocimiento del mundo para reemplazarla por una que sí lo es, ignorando la derivación lógica.

4- Un silogismo en la cuarta figura y modo AO, dos modelos y sin conclusión válida, con la siguientes premisas:





*Todos los médicos son especialistas/ Algunos especialistas no son obstetras*

Dos sujetos respondieron que no hay conclusión válida. El efecto de la figura es determinante en cuanto a la dirección de la conclusión ( A-C). Si se sigue literalmente la formulación de los cuantificadores en las premisas, esta dirección produciría la extraña conclusión " todos los a no son b"; pero la construcción del primer modelo mental que integra ambas premisas no produce esta conclusión; sino " Algunos A no son C".

[ a ]	b	
[a ]	b	[c]
	b	[c ]

Efectivamente predomina la respuesta "algunos A no son C" (Johnson-Laird y Bara: 75%; propio: 53,50 %). También en ambos aparecen respuestas "Algunos A son C" (aproximadamente 20%). Tampoco hay diferencia en cuanto a la respuesta "no hay conclusión válida", lo que significa que en ambas ocasiones pocos sujetos construyeron el segundo modelo que les permitiría advertir que no se deriva nada.

Los resultados propios y los que resultan de la comparación apoyan la hipótesis de que la credibilidad actúa sesgando el proceso inferencial mismo. Los contenidos utilizados son en todos los casos creíbles, por lo que puede suponerse que no influyen en la interpretación de las premisas.

Cuando no actúa el efecto de la figura ( 1 y 3 ) se observa una importante diferencia en el número de respuestas "no hay conclusión válida" que se obtuvieron con contenidos familiares pero neutros, en relación a las obtenidas con contenidos que forman parte del sistema de creencias de los sujetos. En cambio esta diferencia no es importante cuando actúa el efecto de la figura, especialmente en la figura cuarta en la que este efecto parece ser mas decidido.

En general, pocos sujetos demostraron realizar el procesos de construcción y evaluación de modelos alternativos, proceso en el que reside propiamente la competencia deductiva. Tendencia que se ve acentuada por el efecto de las creencias, ya que al obtener una primera conclusión compatible con sus creencias la aceptaron sin intentar ponerla a prueba; lo cual indica que tienden a prescindir de la validez para privilegiar las convicciones.

#### ***b) creencias y modelos mentales alternativos***

El efecto de las creencias se manifestó también en otro tipo de situación, en la que se manipuló el efecto de las creencias en la motivación para continuar la búsqueda de modelos



alternativos que hagan falsa la primera conclusión obtenida.

Es el caso de un silogismo de modo EAO en la tercera figura y con conclusión válida. Es un argumento de muy alta complejidad que no obtiene prácticamente respuestas correctas en diversas experiencias realizadas, incluida la presente. Supone la construcción y evaluación de tres modelos mentales, tiene conclusión diferente a ambas premisas (particular negativa), y la dificultad adicional de que el modelo que permite obtener la conclusión correcta se construye contra el efecto de la figura.

En esta ocasión comparamos el resultado propio con los que informan Johnson-Lair y Bara y también Durand (1998). Utilicé las siguientes premisas:

Ningún médico es sacerdote

Todos los médicos son personas tranquilizadoras.

Algunas personas tranquilizadoras no son sacerdotes (única conclusión válida- C-A)

Por su estructura, figura tercera, para lograr la contigüidad de término medio, después de leída la segunda premisa y advertido que no puede formarse un modelo integrado de ambas, el modelo de la primera premisa es invertido espacialmente para integrar la segunda. O bien se inicia por la segunda premisa invirtiendo el modelo y se relea la primera. La combinación de los modelos mentales de las premisas produce un primer modelo que sostiene la conclusión Ningún A es C, o bien Ningún C es A.

[a]

[a]

[b]] [c

[b]] [c

----

Esta es efectivamente la respuesta que los sujetos producen con mayor frecuencia.

El estudio de Johnson-Laird y Bara registra 60% de respuestas con el cuantificador ninguno, en una u otra dirección; mientras que el de Sonia Durand (1998) registra 67%.

En cambio, solamente uno de mis sujetos produjo esta respuesta. La diferencia es significativa y se debe a la saliencia afectiva del contenido. La conclusión producida por el modelo mental inicial sería "Ningún sacerdote es persona tranquilizadora" o su inversa; son conclusiones no coherentes con convicciones socialmente modeladas.

La relación de disyunción entre sacerdotes y personas tranquilizadoras que señala el modelo no es aceptada y motiva la construcción de un modelo alternativo. Como los "c "



(personas tranquilizadoras) no están exhaustivamente representados respecto a los "a" ( sacerdote) porque no lo están respecto a los "b" ( médicos); nada impide que algún "c" pueda ser también "a".

[a]  
[a]                    [c]  
         [b]]        [c]  
         [b]]        [c]

Este segundo modelo sostiene la conclusión algunos a no son c, pero también algunos a son c. La primera corresponde a "Algunos sacerdotes no son personas tranquilizadoras" y obtuvo un 10%, lo que indica que al hallar esta respuesta compatible con sus creencias la proponen sin evaluación de los dos modelos construidos.

La evaluación de los dos modelos muestra que no hay una conclusión compatible con ambos, por lo que es posible que el razonador llegue a la conclusión de que no se puede derivar nada. Tanto mis resultados como los de Johnson -Laird y Bara obtienen una tasa importante de respuestas "no se deriva nada" ( 43% y 35% respectivamente).

El resto de las respuestas indican que se abandona la tarea lógica y se consigna una conclusión acorde con las convicciones, sin ulterior referencia a los modelos mentales (por ejemplo: "todos los sacerdotes son personas tranquilizadoras" o " médicos y sacerdotes son personas tranquilizadoras").

El problema requiere la construcción de un tercer modelo mental y la evaluación del conjunto de modelos para hallar la conclusión válida, pero prácticamente esta instancia no es alcanzada y sólo excepcionalmente se observa alguna respuesta correcta.

En definitiva, en esta forma silogística, contenidos familiares pero neutros respecto al sistema de creencia de los sujetos producen mayoría de respuestas que corresponden al primer modelo construido; quienes continúan y construyen un modelo alternativo advierten que ambos modelos no son compatibles y concluyen que no se deriva nada. Pero prácticamente ningún sujeto construye el tercer modelo que permitiría, si se acompaña de una correcta evaluación, producir la respuesta válida.

En cambio, con contenidos que producen una conclusión no compatible con las creencia, la respuesta inducida por el primer modelo no es aceptada y esto motiva la búsqueda de un modelo alternativo; o bien el abandono de la tarea lógica para producir una conclusión acorde con las creencias.



### *Conclusiones*

En relación al modo como interaccionan la credibilidad y la validez en la construcción y evaluación de modelos mentales, las situaciones inferenciales analizadas ubican los efectos de la credibilidad en el proceso inferencial mismo. O bien desalentando la construcción de algún modelo alternativo de las premisas, porque la conclusión que admite la lectura del primer modelo mental es creíble. O bien alentando esta construcción por la razón contraria, al no ser la primera conclusión obtenida compatible con el sistema de creencias del sujeto. En este último caso se presentó una situación en la que el efecto de las creencias no se debe a la no adecuación de esta conclusión a la verdad empírica, sino a su no adecuación a una convicción socialmente modelada.

Los resultados apoyan la hipótesis de que el contenido influye en el proceso inferencial en el estadio de construcción y validación de los modelos mentales; y no son coherentes con la hipótesis de la existencia de una competencia deductiva subyacente basada en reglas de algún tipo, puesto que en este caso los errores en la actuación se deberían a la interpretación de las premisas y no ocurrían en el proceso inferencial mismo. Son también indicativos de lo determinante que resultan los conocimientos previos y las convicciones; y consistentes con la predicción de que la tendencia será hacia el procesamiento completo en los increíbles y no en los creíbles. Si se aceptan las afirmaciones de la teoría acerca de la naturaleza de la competencia inferencial y de los procesos que la hacen posible, hay que aceptar también que pocos de nuestros sujetos evidenciaron poner en acción las operaciones de comprobación semántica del argumento. Todas las situaciones inferenciales que se les presentaron permiten volver a interpretar y combinar los modelos de las premisas, pero las respuestas consignadas muestran que esta operación solo excepcionalmente fue realizada; lo cual no implica necesariamente que los sujetos carecen de la habilidad necesaria, pero sí que frente a conclusiones acordes con sus creencias no encuentran motivación para ponerla a prueba.



**Referencias Bibliográficas**

Diez, J. A. y Moulines, C. U. ( 1997). *Fundamentos de Filosofía de la Ciencia*, Ariel, Barcelona.

Durand, S. (1998). *Modelos mentales y prácticas deductivas*, U. N. de Entre Ríos. Tesis de Maestría. Inédita.

Evans, Barston y Pollard (1983). *On the conflict between logic and belief in syllogistic reasoning*. *Memory and Cognition*, 11, 295-306.

Johnson-Laird, P (1982). *El pensamiento como habilidad*. En: Carretero y García Madruga (comp). *Lecturas de psicología del pensamiento*. Alianza. Madrid, (1984)

------(1983) *Mental Models*, Cambridge University Press, Cambridge.

------(1984) *El pensamiento como habilidad*. En: Carretero y García Madruga: *Lecturas de psicología del pensamiento*. Ed. Alianza, Madrid.

------(1987) *Modelos mentales en ciencia cognitiva*. En: Norman, D. A. Paidós. Barcelona.

------(1990) *El ordenador y la mente. Introducción a la ciencia cognitiva*. Paidós. Buenos Aires.

------(1993) *Human and machine thinking*, Hillsdale, New Jersey, Laurence Erlbaum Associates.

Johnson-Laird y Steedman (1978). *The Psychology of Syllogisms*. *Cognitive Psychology*, 10, 63-99.

Johnson-Laird y Bara (1984). *Syllogistic Inferences*. *Cognition*, 16, 1-62.

Johnson-Laird y Byrne, R. M.J. (1991): *Deducción*. En: Ayuso, M. del C. *Razonamiento y racionalidad*. Paidós, Barcelona, 1997.

Oakhill, Garnham y Johnson-Laird (1990). *Belief bias effects in syllogistic reasoning*. En : Gilhooly, Keane, Logie y Erdos, *Lines of thinking: reflections on the psychology of thought*, Chichester, Willey.

Riviére, A.(1991) *Objetos con mente*. Alianza Editorial, Madrid.



Santamaría, C (1995). *Introducción al razonamiento humano*. Aianza. Madrid.

Velasco, J. y García Madruga, J.S. (1997) : *El desarrollo de los procesos meta-lógicos y el razonamiento lógico durante la adolescencia* . *Cognitiva* , 9, 139-158