

EPISTEMOLOGÍA E HISTORIA DE LA CIENCIA

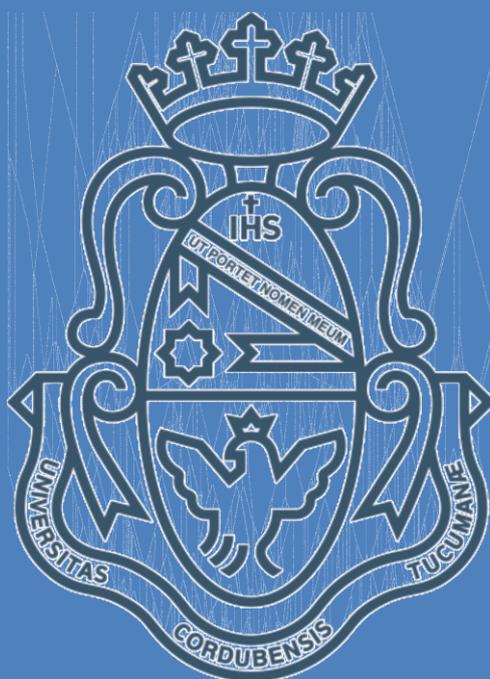
SELECCIÓN DE TRABAJOS DE LAS IX JORNADAS

VOLUMEN 5 (1999), Nº 5

Eduardo Sota

Luis Urtubey

Editores



ÁREA LOGICO-EPISTEMOLÓGICA DE LA ESCUELA DE FILOSOFÍA
CENTRO DE INVESTIGACIONES DE LA FACULTAD DE FILOSOFÍA Y HUMANIDADES
UNIVERSIDAD NACIONAL DE CÓRDOBA



Esta obra está bajo una Licencia Creative Commons atribución NoComercial-SinDerivadas 2.5 Argentina



Resolución de problemas y representaciones visuales¹

José Ahumada / Pío García*

Introducción

En los últimos años se ha incrementado notablemente el interés por el papel que las imágenes tendrían en la solución de problemas. Se ha discutido asimismo su importancia en el contexto del descubrimiento científico y de la creatividad en general. Parte de esta discusión se ha centrado en la posibilidad y, eventualmente, en el modo de diferenciar las representaciones sentenciales de las pictóricas. Pero, el desarrollo de investigaciones sobre razonamiento diagramático y con imágenes no sólo ha generado una discusión sobre cómo caracterizar o evaluar la eficacia este tipo de representación en relación con la representación sentencial, sino también, dentro de la filosofía de la ciencia, ha hecho que se reconsidere la noción de *switch* gestáltico en relación con el cambio conceptual. Filósofos e historiadores de la ciencia como Laudan y Nersessian, no satisfechos con la llamada aproximación lingüística al cambio teórico han argumentado, en varios trabajos, que el cambio conceptual puede entenderse, en términos racionales, desde la perspectiva de resolución de problemas. Pero ambos autores no dicen mucho sobre los mecanismos y razonamientos que tendrían lugar en estos períodos particulares de la historia de la ciencia. La pregunta que surge inmediatamente a partir de la constatación del rol importante de lo visual en el cambio conceptual es si las investigaciones actuales sobre la temática son suficientes para entender esta problemática de un modo distinto a lo que lo entendieron los psicólogos de la *gestalt*. Asimismo, nos preguntaremos si los fracasos y problemas que todavía persisten se deben a una falta de desarrollo de las ideas subyacentes o, si por el contrario, se necesita un modelo totalmente diferente en torno a lo visual para entender cabalmente la resolución de problemas.

El primer paso para comenzar a responder estas preguntas es, obviamente, establecer qué es una representación con imágenes o con diagramas, cómo se diferencia de una representación sentencial y qué efectos tiene este tipo de representación sobre el razonamiento en general.

Qué es una imagen

La pregunta acerca de lo que es una representación visual puede ser reformulada a partir de sus diferencias con las representaciones sentenciales. Se ha argumentado que las representaciones descriptivas contienen información simbólica e interpretada, mientras que las representaciones pictóricas contienen información geométrica y no interpretada o inclusive que la única forma de representación (real) es a través de una cadena de símbolos. Sin embargo, en general, se hipotetiza a favor de la multiplicidad de representaciones, una de las cuales sería la vinculada a imágenes.

Hanson² al hablar de este problema subraya mas bien la distancia entre las representaciones pictóricas y sentenciales con expresiones tales como *abismo* o *laguna*. Aunque da, por cierto, algunas pistas de corte más positivo. En concreto reconoce que los lazos entre

* U.N.C./Secyt.

imágenes suelen hacerse en términos de copia, esto es a través de la *posesión* de ciertas propiedades del original, cosa que no ocurriría en el lenguaje. Elemento este que sirve, en parte, para explicar porqué el lenguaje sería más versátil que las imágenes. Sin embargo problemas como el de la relevancia son vistos, por Hanson, exclusivamente en términos de las representaciones sentenciales, en gran medida a causa de la identificación que hace este autor de lo que aquí llamamos representaciones pictóricas con un lenguaje fenoménico. Hanson sugiere, además, que hay otra diferencia, evidenciada en las distintas funciones y operaciones que pueden realizarse con partes ya sea de una imagen o de un enunciado.

Este parece ser el camino para una fructífera caracterización de las representaciones pictóricas. Así parece entenderlo Scholz,³ quien siguiendo sugerencias de Goodman, ha argumentado a favor de poner el problema en éstos términos.

En ciencias cognitivas la definición más común, y la que ha generado bastante confusión, es que la representación visual es analógica (y la sentencial digital). El problema principal que genera esta definición es que podemos representar imágenes usando una representación sentencial (por ejemplo, podríamos usando letras representar la imagen de una persona), cada símbolo individual (letras) no es en sí mismo analógicos pero tomados como un todo representan una imagen. Así, si preguntamos ¿qué es una imagen?, entonces puede que intentemos responder a partir de las propiedades que hacen que una imagen sea lo que es. Pero esta aproximación nos lleva a un camino sin salida, en tanto no parece haber elementos de este tipo que nos permitan realizar una distinción. No hace falta elegir ejemplos demasiado recalcitrantes para mostrar esto. Las diferentes posibilidades de representación de un cartel o un aviso publicitario en una calle (puede ser visto como una representación pictórica o como una sentencial) pueden darnos una pista de las dificultades que enfrentamos. Scholz sugiere la plausibilidad del siguiente enunciado, que parece ser un buen punto de partida para nuestra discusión: *nada es en sí mismo un símbolo de un tipo específico.*

Este enunciado intenta evidenciar lo que implícitamente afirmábamos en el párrafo anterior: no parece haber características intrínsecas que hagan a un símbolo de determinado tipo. La pregunta, entonces, se desplaza desde *qué es una imagen a cuándo es imagen* o, mejor, bajo qué circunstancias un símbolo es de determinado tipo. Esta clase de aproximación es llamada por Scholz el punto de vista *funcional*. Con esto tendríamos una primera aproximación a una respuesta a la pregunta que nos hacíamos por medio de Hanson, acerca de cuál sería la diferencia entre las representaciones pictóricas y las sentenciales. Este punto de vista funcional, como lo sugiere el mismo Scholz, puede acercarnos a una noción sintáctica. Además este autor sugiere que las imágenes, a diferencia de una representación sentencial, son sintácticamente *densas*. Lo que significa que dado un esquema con infinitos caracteres si estos están ordenados de tal forma que entre dos cualesquiera siempre hay un tercero, entonces dicho esquema es denominado denso. Este tipo de aproximación captura la noción de analógico que discutíamos más arriba sin contradecir el principio que proponíamos, a saber que nada en sí mismo es un símbolo de determinada clase.

Aunque este autor no nos ofrezca una caracterización positiva acerca de cuáles serían las reglas a considerar, de hecho hay ya aquí una respuesta al menos general a este problema: la información visual más que tener algún elemento nuevo, lo que la caracteriza es tener una *organización* distinta de dicha información. Como veremos algunos de los elementos de esta discusión han tenido lugar también en el ámbito de las ciencias cognitivas y de la resolución de problemas en inteligencia artificial.

Sin embargo hay autores que han subrayado el aspecto inferencial para hablar de representaciones pictóricas. Así, para Lindsay lo propio de la representación pictórica está en los mecanismos de inferencia que soporta. En particular afirma que la manera en que están construidas las representaciones pictóricas restringe, y en este sentido posibilita, la realización de inferencias no deductivas.⁴

El problema de las representaciones

La atención dedicada a la pregunta acerca de lo que son las imágenes es una muestra de la importancia que algunos autores le conceden al tema de la representación pictórica. La razón para otorgarle tanta relevancia al problema de la representación en general y a la pictórica en particular, puede formularse de la siguiente manera: “*dos formalismos con similares capacidades expresivas pueden diferir en cuanto a su poder heurístico... facilitando la búsqueda de soluciones.*”⁵ Esta es una de las principales intuiciones que motiva a los investigadores en este campo para la exploración de distintas representaciones. Sin embargo esta afirmación merece un análisis más detallado. Al nivel de *implementación* habría, en general, dos maneras de lograr mayor poder heurístico: restringiendo la búsqueda en el espacio o a través de un ordenamiento (indexación) que permita elegir buenos “candidatos”.⁶ Ésta última aparece como la promesa propia de las representaciones pictóricas. Debemos, sin embargo, recordar una importante distinción. Simon⁷ ha argumentado a favor de que los razonamientos visuales pueden ser representados usando diagramas y que, a su vez, los diagramas pueden ser llevados a un sistema simbólico. El diagrama es, en términos generales y tal como lo entiende Simon, una *simplificación* de lo que describe, eliminando muchos factores contextuales que pueden llegar a ser relevantes para relacionar hipótesis o imágenes. Así, por ejemplo, las representaciones visuales parecen permitir la atribución de relaciones causales, tal como lo sugiere Thagard,⁸ restringiendo la búsqueda de dichas relaciones y por ende acotando el problema de la relevancia. Sin embargo hay una dificultad asociada con este concepto. Generalmente las referencias al poder heurístico, que señalábamos en el párrafo anterior, son descriptas en términos de *eficiencia*. No obstante, un problema a tener en cuenta es que en este contexto no es claro hablar *simplemente* de mayor eficiencia. Podemos entender esta cuestión en referencia a máquinas artificiales, pero es todavía bastante difuso el problema de cómo han solucionado este punto los seres humanos. Lo que queremos decir, en otras palabras, es que puede que la eficiencia de una representación dada no dependa *sólo* de la *máquina virtual*⁹ sobre la que está implementada, sino del hardware que la soporta.¹⁰

Si bien una representación pictórica o espacial puede tener mayor poder heurístico (en algún sentido ser más eficiente), esto no significa que sea útil en *cualquier contexto*. Así y todo este tipo de representaciones tiene un apreciable campo de aplicación. Por ejemplo, el uso de representaciones espaciales puede ser conveniente en ciertos *contextos* aunque, de suyo, la representación *propia* de ese campo sea otra. Con frecuencia, en el contexto de aprendizaje de los números y de sus operaciones los maestros apelan a imágenes y relaciones espaciales, a veces en forma de operaciones sobre listas. Los lenguajes basados en listas son utilizados a partir de sus propiedades espaciales, a pesar de que un tratamiento lógico es posible, en tanto el primero pareciera más intuitivo y directo. Es más, muchas personas para la solución más eficiente de algunas operaciones matemáticas simples recurren al uso de imágenes y relaciones espaciales. Esto nos da la pista de que para resolver algún problema,

muchas veces apelamos a diferentes representaciones. Es así que, aparentemente y como ya lo hemos sugerido en anteriores trabajos, una habilidad importante en las personas, en orden a lograr una mayor eficiencia en la solución de problemas, es la de poder cambiar de representación.

Es más, Barwise y Etchemendy creadores del programa Hyperproof sostienen que hay teoremas que pueden ser demostrados usando representaciones heterogéneas pero que no pueden ser demostrado usando solamente una representación textual o diagramática. Para estos autores la ventaja de la representación diagramática reside en la posibilidad de hacer explícita la información implícita de modo tal que puede llegar a sustituir la inferencia. Varias objeciones de peso se le han formulado a este proyecto, la primera es que el uso de diagramas solo tiene ventaja respecto a la representación textual para determinada cantidad de información. La segunda, es la dificultad de representar información parcial que obliga a la incorporación de convenciones llevando a que la inferencia no sea propiamente reemplazada por la observación sino por búsqueda de observación más la habilidad de juzgar lo que es relevante y lo que es irrelevante. La distinción máquinas virtuales/reales también mostrarían que las ventajas del Hyperproof son sólo pragmáticas.

Este tipo de elementos dio pie a la búsqueda de una manera de representar computacionalmente la información de tipo visual. Entre los desarrollos actuales veamos brevemente el correspondiente a de Janice Glasgow, en tanto, por un lado puede ser una buena muestra de lo que se ha avanzando en este campo y de los problemas que todavía permanecen.

Representaciones visuales en Inteligencia Artificial: la propuesta de Glasgow
Típicamente una representación dada o una familia de representaciones expresará cierta información de manera explícita, mientras que otra información permanecerá implícita. De aquí que el siguiente paso sea la implementación de algún medio para explicitar esta información. Los mecanismos que realizan esta tarea son llamados por algunos autores inferencias (así lo entienden, entre otros, Barwise¹¹ y Lindsay¹²). En este trabajo nos concentraremos en la primera parte de esta tarea, dejando el problema de los mecanismos inferenciales para una elucidación posterior. Además del problema general de las representaciones nos concentraremos en las cuestiones típicas de las representaciones computacionales. Esta aclaración es importante en tanto reconocemos que éste es un ámbito más restringido que el de las representaciones en general y con problemáticas más acotadas, tales como el de su implementación.

A pesar de que es un lugar común, dentro del campo de la inteligencia artificial, hablar de la falta de una teoría general de la representación del conocimiento, siguiendo a Lindsay¹³ podemos, al menos, plantear ciertos requerimientos mínimos que, intuitivamente, ésta debiera cumplir.¹⁴ En concreto no bastaría con presentar la *estructura* de dicha representación sino que se requiere especificar *cómo es construida* y *cómo puede ser accedida*. Un mismo ordenamiento de datos puede ser utilizado de diferentes maneras, lo cual implica, a veces, la conformación de distintas representaciones. Esto nos pone en la tesitura de *exigir* reglas de acceso y manipulación de datos como parte de una representación implementada.

El trabajo de Glasgow, además de cumplir con estos requisitos mínimos, plantea de manera original este problema. Probablemente su mayor mérito sea tomar algunas sugerencias de la psicología cognitiva y proponer una subdivisión de las representaciones visuales. Así,

esta autora, pretende definir un esquema de conocimiento representacional que permita poner en primer plano las propiedades más importantes, tanto *espaciales* como *visuales*, de una *imagen*.¹⁵ Por un lado experimentos cognitivos parecen avalar la existencia de una memoria visual distinta de una memoria verbal y que en el aspecto de recuerdo parece, la primera, superior a la segunda. Esto plantea una primera separación entre representaciones visuales y descriptivas. Los experimentos llevados adelante por Kosslyn parecen sugerir, además, que lo que se *conserva* en este tipo de memoria son las relaciones espaciales, los tamaños relativos y las distancias relativas. Pero además se descubrió que hay pacientes que tienen dificultades para *localizar* objetos, a pesar de que pueden *reconocerlo*. Y otros pacientes muestran patrones visuales opuestos: pueden reconocer objetos pero no localizarlos.¹⁶ Pese a que estos datos sirven como sugerencia para el esquema representacional de Glasgow, esta autora está más preocupada por la expresividad y eficiencia de la implementación computacional que por imitar el cerebro o la mente humana. No obstante lo cual puede notarse, en el sistema de Glasgow, la influencia de otros elementos de la psicología cognitiva tradicional.

La sección representacional del sistema está dividida en tres partes. En primer lugar tenemos una **representación profunda** la cual es llamada, sugestivamente, memoria de largo plazo. Ésta es la encargada de guardar información que será utilizada por la "memoria de trabajo". Esta representación profunda es codificada usando los *frames* de Minsky, los cuales teniendo una estructura jerárquica, permiten, entre otras cosas, utilizar la importante propiedad de *herencia*.¹⁷ Evidentemente la propiedad de herencia va a ser empleada en la construcción de imágenes a partir de sus componentes. Esta solución no es original de Glasgow y podemos rastrear su génesis en los trabajos acerca de "análisis de características" de Biederman a fines de los ochenta.¹⁸ Un ejemplo simple de cómo funciona este esquema es la construcción del mapa de Europa a partir de sus países integrantes. La intención es aplicarlo a contextos más complejos, en tareas que requieran la utilización de grandes bases de datos.¹⁹

Además de la **representación profunda**, tenemos la **representación visual y la espacial**. La representación visual corresponde al componente visual de la imagen, no solo en términos de información geométrica sino en elementos tales como volumen, brillantez, textura, etc.²⁰ Esta información es básicamente no interpretada. En cambio la correspondiente al componente espacial de las imágenes necesita hacer referencia a relaciones espaciales entre partes significativas de dicha imagen. Con este fin se utiliza un tipo de estructura²¹ de datos que preserva las propiedades espaciales y las relaciones topológicas de las imágenes.

Conclusiones

Es evidente que se ha recorrido un camino en la tarea de obtener una representación (computacional) pictórica. De la investigación actual puede verse que son variadas y diferentes las vías de acceso elegidas para atacar esta cuestión. No obstante, creemos que existe un núcleo de problemas que la mayoría de estas corrientes de investigación comparte.

Sistemas como el de Glasgow parecen desarrollar algunas de las ideas que generalmente se sostienen acerca de lo que debería ser una representación de este tipo. En particular algunas de las sugerencias de Scholz, como su punto de vista funcional, parecen encontrar eco en el programa de Glasgow. La división entre una representación visual y una espacial,

si bien, en principio, de aplicación restringida, parece fértil en tanto permite diferenciar algunos de los atributos que aparecen como propios de una imagen. Glasgow sugiere que sistemas como éste pueden ser aplicados en ámbitos donde la tarea sea la clasificación o el reconocimiento de información visual. Pero una atenta inspección a las funciones que se utilizan en cada uno de los módulos representacionales evidencia que, probablemente, el trabajo mayor por hacer es en este ámbito.²² Decíamos arriba que una representación visual permitiría, aparentemente, un acceso más fácil y rápido a la información, pero subrayábamos, asimismo, su carácter acotado y específico. Los diagramas y las imágenes, en tanto han sido descriptos como dos aproximaciones, con ciertas diferencias, al problema de la información y su manejo, exigirían, en apariencia, la especificación de funciones o reglas heurísticas diferentes.

No obstante este tipo de intento parece, todavía, un poco lejano de contextos más complejos como el de descubrimiento científico. Así, si bien es casi una constante sindicarse la capacidad heurística, entendida en términos de relevancia o de simplicidad, como la virtud principal de las imágenes y los diagramas, sin embargo no es todavía claro de que manera explotar computacionalmente su potencial.

En este sentido, Nancy Nersessian²³ ha hecho esfuerzos para ofrecernos un relato, en términos cognitivos pero con especial atención a los casos históricos de cambio conceptual, del uso de diagramas e imágenes. Ha notado algo que generalmente ha pasado inadvertido, la conexión entre psicología y epistemología depende, evidentemente, de la teoría psicológica que uno adopte, y la psicología de Gestalt ha constituido un obstáculo que estas investigaciones sobre imágenes, desde la perspectiva que estamos abordando, habría comenzado a remover.

El inconveniente que presenta su concepción es que no determina ni sugiere cuáles serían los mecanismos que operan en los procesos de cambios conceptuales. Para Nersessian las imágenes son importantes porque permiten lograr una interpretación de la información sentencial y porque ofrecería claras ventajas heurísticas. Su compromiso con la gradualidad del cambio y con la aproximación clásica a la resolución de problemas le impide, en principio, tomar una posición no ortodoxa respecto al razonamiento con imágenes, como sería otorgarle ventajas epistémicas por sobre el razonamiento sentencial. Así parece, por un lado, haber líneas de investigación que muestran un avance en la caracterización de las representaciones pictóricas, haciendo hincapié en el desarrollo de las implementaciones computacionales. Y por otro lado hay investigaciones cognitivas de los casos históricos de descubrimientos que involucraban el uso de imágenes. Pero todavía no se ha llegado a la convergencia de las características aparentemente complementarias de estas exploraciones. El desafío de aquellos que se proponen aprovechar las virtudes de las representaciones visuales y sus posibilidades inferenciales dependen, al parecer, de la posibilidad de esta convergencia.

Notas

¹ Este trabajo ha sido financiado a través de un subsidio de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Córdoba y CONICOR, dirigido por el prof. Víctor Rodríguez.

² Hanson, N.R. [1958].

³ Scholz, Oliver. [1993].

⁴ Ejemplos de este tipo de aproximación son las inferencias que se pueden realizar a partir de la definición de relaciones espaciales simples como arriba-abajo o izquierda-derecha. (Cfr. Linsay, R. [1995].)

- ⁵ Sloman, [1995]. Pág. 10.
- ⁶ Sloman, A. [1995]. Pág. 12.
- ⁷ Larkin, J.H. & Simon, H.A. [1987]
- ⁸ Thagard, P. and Shelley, C. [1997]
- ⁹ El concepto de máquina virtual es utilizado para postular la independencia relativa de una representación respecto de la base física o de otras representaciones. Las razones de esta independencia estarían en la definición de una ontología (estructuras abstractas permitidas) y conducta (operaciones sobre el conjunto de datos). Desde esta perspectiva la implementación de lenguajes de alto nivel como el Lisp o el C pueden ser vistos como *máquinas virtuales*.
- ¹⁰ En nuestro Ahumada, J. y García, P. [1998], habíamos propuesto esta hipótesis. Además, esto nos pone en alerta sobre la idea de que siempre es mejor imitar la manera humana de representar problemas
- ¹¹ Barwise and Etchemendy [1995]
- ¹² Lindsay, R [1995].
- ¹³ Lindsay, R. [1995]
- ¹⁴ Los requerimientos que plantea Sloman [1995] son más generales (sintaxis, pragmática, semántica y estrategias de inferencias) pero a los fines de la implementación estos parecen suficientes.
- ¹⁵ Glasgow, J & Papadias, D.[1995].
- ¹⁶ Glasgow, J & Papadias, D.[1995] pág. 439. Es interesante notar, asimismo, que pueden eventualmente reconocer objetos por el *tacto* o a través de un *sonido* característico.
- ¹⁷ Aquí se almacena tanto información *proposicional* como *procedural*
- ¹⁸ Algunos de elementos de esta historia se encuentran en Ahumada, J. y García, P. [1998].
- ¹⁹ Un ejemplo más complejo es la determinación de la estructura de una proteína en términos de un cristal, tarea que puede ser vista, de acuerdo con Glasgow, como la de construir un agregado, en varias capas, de una serie de subestructuras a partir de una base de datos. El problema en sí es similar al del mapa, solo que siendo de una complejidad mayor muestra la eventual utilidad del mecanismo propuesto.
- ²⁰ Es construida a partir de la representación profunda y su implementación computacional es a través de "arreglos", un tipo de estructura que, en general, permite almacenar elementos y acceder a ellos por medio de un índice.
- ²¹ Un arreglo simbólico con la característica especial de que cada elemento de este array denota alguna parte significativa de la imagen. Este tipo de arreglo es denominado arreglo simbólico anidado, el cual permite implementar un orden jerárquico.
- ²² Las funciones son solo las habituales para manipular imágenes y relaciones espaciales simples.
- ²³ Nersessian, N. [1995].

Bibliografía

- Ahumada, J. y García, P. "Descubrimiento científico y búsqueda de patrones" en *Episteme*, v.3, n.7 1998.
- Barwise and Etchemendy. "Heterogeneous Logic" en *Diagrammatic Reasoning* Edits. Glasgow, Narayanan and Chandrasekaran. AAAI Press/ MIT Press 1995
- Glasgow, J & Papadias, D. "Computational Imagery" en *Diagrammatic Reasoning* Edits. Glasgow, Narayanan and Chandrasekaran. AAAI Press/ MIT Press 1995.
- Hanson, N.R. *Patterns of Discovery*. Cambridge University Press. 1958.
- Larkin, J.H. & Simon, H.A. (1987) "Why a diagram is (sometimes) worth 10,000 words", *Cognitive Science*, vol. 11, pp. 65-100.
- Lindsay, R. "Images and Inference" en *Diagrammatic Reasoning* Edits. Glasgow, Narayanan and Chandrasekaran. AAAI Press/ MIT Press 1995
- Nersessian, N. "How do Scientists Think?. Capturing the Dynamics of Conceptual Change in Science" en *Diagrammatic Reasoning* Edits. Glasgow, Narayanan and Chandrasekaran. AAAI Press/ MIT Press 1995
- Scholz, Oliver. "When is a Picture?" *Synthese* 95: 95-106, 1993.

- Sloman, A. "Towards a general theory of representations" en *Forms of Representation*, Peterson, D. Ed. Intellect Press, 1995.
- Sloman, A. "Varieties of Formalisms for Knowledge Representation" [commentary on "The imagery debate revisited: a computational perspective" by Janice Glasgow] in: *Computational Intelligence*. Special issue on Computational Imagery, Vol. 9, No. 4, November 1993
- Sloman, A. "Musings on the Roles of Logical and Non-logical Representations in Intelligence" en *Diagrammatic Reasoning*, Edits. Glasgow, Narayanan and Chandrasekaran. AAAI Press/ MIT Press 1995.
- Thagard, P., and Shelley, C. P. (1997). "Abductive reasoning: Logic, visual thinking, and coherence". In M.-L. Dalla Chiara et al. (Eds.), *Logic and scientific methods*. Dordrecht: Kluwer, 413-427.