

Buscando a Hal desesperadamente: de la psicología cognitiva a la psicología del conocimiento*

Juan Ignacio Pozo
Universidad Autónoma de Madrid

A pesar de su predominio en los últimos cincuenta años y de sus indiscutibles aportaciones teóricas y empíricas, la psicología cognitiva no ha podido cumplir algunas de sus promesas esenciales para dar cuenta del funcionamiento de la mente humana. En este artículo se destaca que la mayor parte de los componentes cognitivos que no han podido abordarse desde el enfoque computacional dominante son precisamente los que diferencian a la mente humana de otros sistemas cognitivos (las emociones, las intenciones, el aprendizaje conceptual, el significado), que aquí se resumen en la capacidad de la mente humana para manejar conocimientos, definidos como representaciones explícitas, y no sólo, como el resto de los sistemas cognitivos, representaciones de naturaleza implícita. Estudiar esas representaciones explícitas requiere desarrollar, en el marco de la psicología cognitiva, una psicología del conocimiento que recupere para la investigación psicológica tres componentes esenciales de la mente humana olvidados por la clásica psicología cognitiva: los contenidos, la conciencia y la cultura. Sólo así será posible que Hal, el robot «humano» ideado por Arthur C. Clarke, pueda ser comprendido desde la psicología cognitiva.

Palabras clave: representaciones explícitas, representaciones implícitas, conocimiento, contenido, conciencia, cultura

After five decades of research, the prevailing computational approach in cognitive psychology has made many important contributions, both theoretical and empirical, to our understanding of how the human mind

* Este trabajo ha contado con el apoyo de la DGESIC a través del proyecto PB98-095. Quiero agradecer los comentarios y sugerencias de María del Puy Pérez Echevarría, Mar Mateos y Nora Scheuer a borradores previos de este trabajo. Parte de las reflexiones aquí expresadas las hemos compartido en el marco del curso de doctorado «De la psicología cognitiva a la psicología del conocimiento», que impartimos conjuntamente. Igualmente quiero agradecer las críticas recibidas a través de los revisores anónimos, que me han ayudado a perfilar mejor algunos de los argumentos desarrollados, además de abrir nuevas preguntas para algunas de las cuales, he de reconocer, no tengo respuesta.

Correspondencia: Departamento de Psicología Básica. Facultad de Psicología. Universidad Autónoma de Madrid. Campus de Cantoblanco. 28049 Madrid. Correo electrónico: nacho.pozo@uam.es

works. However, some of its most significant promises are still unfulfilled; paradoxically, these promises are related to the more specifically human components of the human mind such as emotions, intentions, conceptual learning, and meaning. We suggest that all these components can be explained in terms of the specific (in the strong sense) capacity of the human mind, when compared with other cognitive systems, to work with knowledge. To study how the mind knows, or uses explicit representations and not only implicit ones (as do the remaining cognitive systems) we must assume, within the cognitive framework a "psychology of knowledge" approach based on the study of three specific dimensions of human knowledge – content, consciousness and culture – that are clearly missing in traditional cognitive research. Only through this approach we will be able to make sense of Hal, the «human» robot that Arthur C. Clarke once imagined.

Key words: Explicit representations, implicit representations, knowledge, content, consciousness, culture.

Introducción: buscando a Hal

Una vez traspasada la frontera simbólica del año 2001, y a pesar del sueño de Arthur Clarke, que Stanley Kubrick convirtiera en imágenes a ritmo de vals, la psicología cognitiva sigue buscando desesperadamente a Hal entre sus sistemas de cómputo o más recientemente entre sus redes neuronales y sus sistemas conexionistas. ¿Por qué Hal no vive ya entre nosotros? ¿Qué le falta a la psicología cognitiva, y con ella a la ciencia cognitiva, para integrar en sus sistemas de cómputo, y en su caso similar, componentes tan básicos, y tan específicos (es decir propios de nuestra especie), como las emociones, las intenciones, el significado del discurso o el aprendizaje conceptual? Tal como señalara Manuel de Vega (1998), el indudable éxito de la psicología cognitiva, su empuje arrollador en las últimas décadas, su ingente aportación empírica y teórica, no le ha permitido sin embargo cumplir el sueño de encontrar a Hal. Debemos admitir, cumplidos ya los primeros años después de Hal, que el sujeto de la psicología cognitiva, del que nos hablara Ángel Rivière (1987), con toda su complejidad y sofisticación computacional, no es aún un sujeto plenamente humano.

En los últimos años la psicología cognitiva ha ido tomando conciencia de los límites del enfoque cognitivo clásico, basado en el procesamiento de información y el cómputo digital de representaciones, algunos de los cuales hacía tiempo que venían siendo anunciados (p. ej., Bruner, 1983; Caparrós y Gabucio, 1986; Carretero, 1997; de Vega, 1981, 1982; Rivière, 1987, 1991). Por fortuna han ido surgiendo nuevas alternativas teóricas con el fin de superar algunos de estos límites. Buena parte de estas alternativas, quizás las que más han influido en la nueva orientación teórica de la psicología cognitiva, han asumido que el nivel representacional, en el que se situaba la investigación cognitiva, era demasiado molar o global y que se requería adoptar unidades de análisis más reducidas, de forma que esas representaciones no serían sino el efecto de la conexión masiva de unidades de cómputo, entre las que esas representaciones es-

tarían de hecho distribuidas. El enfoque conexionista surge así como alternativa a los modelos clásicos de «computaciones sobre representaciones» (Rivière, 1991), apoyado además por el desarrollo de nuevas técnicas de neuroimagen que han permitido un notable desarrollo de las neurociencias que, más allá de la simulación, permiten el estudio empírico de esas conexiones en sistemas neuronales reales.

¿Estará Hal oculto en los pliegues computacionales de esas redes neuronales? ¿Será únicamente el efecto de la conexión masiva entre esas unidades de información, de la misma forma que, supuestamente, lo son también las representaciones cognitivas? Si el conexionismo ha podido dar cuenta, en su nuevo nivel de análisis, de algunos de los fenómenos representacionales clásicos e incluso ha podido abordar situaciones no explicables en términos de representaciones simbólicas (ver Clark, 1997), ¿por qué aún no ha podido encontrar a Hal? Tal vez la razón de que aún no se le haya encontrado en los sistemas de cómputo de la psicología cognitiva, y por tanto sigamos buscándolo desesperadamente, es simplemente que no se encuentre ahí, o al menos que no se encuentre *sólo* ahí. Aunque sin duda Hal tiene que responder, de algún modo aún inexplicable, a la lógica no sólo estadística sino sobre todo bioquímica de esas conexiones, también debe ser causa de los cambios que se produzcan en esos circuitos, debe tener un efecto físico real en ellos (Cairns-Smith, 1996; Lewontin, 2000), por lo que su funcionamiento como sistema no podría reducirse a los principios que rigen esas conexiones, ni tampoco a las representaciones manipuladas por el cognitivismo clásico. Si para el conexionismo, el nivel representacional es demasiado molar, desde otros enfoques igualmente cognitivos, se considera que ese nivel representacional es en cambio demasiado molecular o reducido (Pozo, 1997), y que para abordar los problemas cognitivos planteados por Hal es necesario ir más allá de las representaciones. Es en este sentido en el que la psicología cognitiva, en su nivel de análisis representacional, debe dar paso a una psicología del *conocimiento* para hacerse *humana/mente* reconocible (Pozo, 2001) sin por ello renunciar a algunas de sus aportaciones más sustanciales.

Sin embargo, pasar del nivel estrictamente cognitivo, el de las representaciones, ya sean simbólicas o distribuidas entre unidades, al estudio del conocimiento supone, en mi opinión, trascender algunos de esos límites clásicos de la psicología cognitiva, que ha definido un sujeto individual, socialmente aislado, dotado de recursos formales de procesamiento pero carente de contenidos y de conocimiento sobre sus propias representaciones y procesos. Estas carencias definirían tres componentes (cultura, contenidos y conciencia) que sería necesario incorporar para que Hal fuera no sólo un autómatas cognitivo sino un sujeto con conocimiento. Tal vez sea en la comprensión de estas limitaciones, y en su posible superación, donde debamos buscar a Hal. Para ello, sin abandonar el enfoque cognitivo general, al menos entendido en un sentido amplio (De Vega, 1998; Rivière, 1987), debemos proponernos ir más allá de esas restricciones impuestas por la corriente dominante del procesamiento de información que, en cualquiera de sus variantes, asume que la mente humana puede reducirse a un sistema de cómputo.

¿Puede ser Hal sólo un sistema de cómputo?

Entendida en un sentido amplio, la psicología cognitiva se ha caracterizado por definir como objeto de estudio las representaciones y los procesos mediante los que se manipulan esas representaciones (Rivière, 1987). De esta forma, la gran aportación de la psicología cognitiva, o más en general de las ciencias cognitivas, ha sido definir un nivel de explicación intermedio a los estados mentales y a la conducta, que hasta la llamada revolución cognitiva habían sido los niveles de análisis desde los que se había intentado elaborar una psicología científica. Según el acercamiento cognitivo, lo que las personas somos y hacemos, lo que queremos y pensamos, depende no tanto de nuestras conductas eficientes, como suponía el conductismo, o de los estados mentales que acompañan a nuestras conductas, que eran el objeto de estudio de los enfoques introspectivos o fenomenológicos, como de las representaciones que tenemos con respecto al mundo y a nosotros mismos. Entendida así la psicología cognitiva, en ella cabe no sólo el enfoque dominante del procesamiento de información, sino también la aportación de autores como Piaget, Vygostki o incluso la Gestalt (Carretero, 1997; Pozo, 1989; Rivière, 1987). Sin embargo, si bien todas estas aportaciones son, en el sentido antes señalado, genuinamente cognitivas, ya que sitúan el nivel explicativo en el marco de sistemas de representación, aunque sean de diferente naturaleza, aceptando la definición de la psicología cognitiva como una categoría natural al modo de Rosch, tal como nos propuso Rivière (1987, 1991), hay que admitir que algunas de estas posiciones serían «más cognitivas» que otras, en el mismo sentido en el que las personas tendemos a creer que algunos animales son «más mamíferos» que otros.

De hecho, las posiciones que se acercarían más al prototipo de la psicología cognitiva son aquellas que han asumido que esas representaciones pueden ser traducidas, y en realidad reducidas, a cálculos formales. El enfoque dominante, el que en efecto ha hecho las contribuciones más decisivas que permiten hablar propiamente de la psicología cognitiva como un ámbito pujante de investigación, ha sido el procesamiento de información, según el cual la mente humana puede entenderse como un conjunto de sistemas de cómputo, o más recientemente de unidades de información, cuyo producto serían las representaciones. Hay por supuesto otras aportaciones cognitivas ajenas a la tradición computacional dominante, como por ejemplo, las de Piaget, Vygostki o la reciente psicología de la instrucción (p. ej., Carretero *et al.*, 1991; Voss, Wiley y Carretero, 1995; Reigeluth, 1999), que no serán objeto de análisis en este trabajo, entre otras cosas porque parte de los argumentos aquí desarrollados a favor de una psicología del conocimiento recuperan algunas de las aportaciones de esos otros enfoques cognitivos, que podríamos considerar alternativos, que han estado en buena medida dedicados a estudiar al sujeto de conocimiento más que al sujeto cognitivo, tal como lo entiende el enfoque computacional. Esa psicología cognitiva dominante, a partir de la analogía establecida por Shannon entre el funcionamiento de un interruptor electrónico y la lógica binaria, supuso de hecho reducir las representaciones a mera información, es decir, a cálculos estadísticos. Todo lo que se necesitaba para representar mundos complejos y actuar

en ellos era un potente sistema de cómputo que permitiera establecer reglas formales que predijeran la conducta de ese mundo y las interacciones con él (enfoque de computaciones sobre representaciones) o establecer pautas de conexión probabilísticas entre unidades de información suficientemente dinámicas para representar los cambios que tienen lugar en el mundo, incluyendo al propio sistema cognitivo como parte de ese mundo (enfoque conexionista).

Esta reducción de las representaciones a mera información implica que todos los mecanismos psicológicos pueden también explicarse en términos de cálculos probabilísticos, si entendemos la información en un sentido estricto («el número de opciones que tenemos al tratar con una serie de ítems», según la precisa definición de Reber, 1995, p. 369, en su *Dictionary of Psychology*). Es bien cierto que en el enfoque simbólico —o de computaciones sobre representaciones (Rivière, 1987)— el sistema cognitivo humano se concebía como un «manipulador de símbolos» más que como una máquina de Turing que elaborase cálculos estadísticos. Pero también es cierto que esos símbolos, debido a su carácter intrínsecamente informacional, es decir a su carencia de contenidos (p. ej., Pozo, 2001; Shanon, 1993), tenían una naturaleza amodal y arbitraria (Barsalou, 1999), propia de la lógica matemática en la que se basa el funcionamiento de todo sistema estrictamente informativo. Este intento de reducir el contenido a forma, o la representación a información, aparentemente desdibujado en el procesamiento de información simbólica, queda de manifiesto no sólo en los recientes desarrollos conexionistas, que en mi opinión profundizan en la propia esencia computacional de los modelos cognitivos, convirtiendo la en ocasiones vaga metáfora computacional en mecanismos eficientes de cómputo, sino también en los mecanismos de adquisición o cambio de representaciones propuestos por la psicología cognitiva del procesamiento de información, que también en la tradición clásica o simbólica se reducen esencialmente a mecanismos de detección de contingencias o regularidades en el ambiente (Pozo, 1989).

A pesar de los problemas que este reduccionismo ha traído consigo, entre otros como veremos expulsar a Hal del paraíso cognitivo, las contribuciones realizadas desde este enfoque computacional han sido tan abrumadoras, tanto cuantitativa como cualitativamente, que durante varias décadas se trató sin duda de un programa de investigación progresivo, en el sentido de Lakatos, por lo que resultaba muy difícil poner en duda sus planteamientos. De hecho, hemos de aceptar con de Vega (1998, p. 34) que «en las tres últimas décadas hemos avanzado en el conocimiento de los procesos mentales más que en los tres últimos siglos». Y ello ha sido posible gracias a entender la mente y la conducta humanas en términos de representaciones regidas por reglas formales de cómputo.

De hecho, a partir de la llamada revolución cognitiva, la investigación científica dispone de marcos teóricos y recursos metodológicos para enfrentarse a un nuevo tipo de sistemas u objetos de estudio, más allá de los tradicionales sistemas *físicos* (o de energía) y de los sistemas *biológicos* (dotados de información genética): los sistemas *cognitivos* (que manejan representaciones). En este sentido, la revolución cognitiva, más allá de sus efectos inmediatos sobre la psicología, ha supuesto la segunda revolución mecanicista en la historia de la ciencia (Rivière, 1991), de forma que ha permitido completar el viejo sueño de hacer

una ciencia mecanicista que incluya no sólo los objetos materiales (física) sino también los organismos (biología) e incluso los sistemas mentales (ciencia cognitiva) (Pozo, 2001). Pero a pesar de su gran éxito a corto plazo y de sus innegables promesas, esta revolución no ha logrado completar del todo ese sueño y ha traído consigo también notables frustraciones (Bruner, 1990). Mientras que las otras dos grandes revoluciones digitales (o computacionales) producidas a mediados del siglo pasado, la revolución genética y cibernética, han tenido un éxito esplendoroso, que promete cambiar radicalmente nuestra forma de vida, las consecuencias sociales y tecnológicas de la revolución cognitiva son mucho más modestas y muchas de sus promesas han quedado incumplidas. Y entre ellas, en un lugar destacado está nuestro Hal, la eterna promesa de la ciencia cognitiva, el sueño incompleto del mecanicismo¹.

Aunque cabrían otras explicaciones para esta ausencia de Hal (ver por ejemplo, De Vega, 1998), desde mi punto de vista la principal razón es que Hal (ese Hal que todos nosotros llevamos dentro, un objeto no sólo con mente, según dijera Rivière, 1991, sino también con conocimiento del mundo y de sí mismo), siendo como sin duda es un sistema de cómputo, *no puede ser sólo un sistema de cómputo*. Obviamente, la psicología cognitiva debe ayudarnos a comprender los mecanismos automáticos, los procesos implícitos mediante los que realizamos gran parte de nuestros cómputos, procesos que sin duda debemos realizar Hal y todos nosotros para generar la mayor parte de nuestras representaciones. Pero muchos de esos procesos implícitos no son específicamente nuestros, no diferencian a la cognición humana (o *haliana*) del funcionamiento de otros sistemas de cómputo, como el programa de ordenador desde el que escribo, el sistema experto que regula el vuelo de un avión, una máquina que juega al ajedrez o realiza diagnósticos clínicos especializados, o, para el caso, del funcionamiento intelectual de los primates superiores. ¿Qué es lo que hace que Hal sea Hal? ¿Qué es lo que diferencia a la mente humana de otros sistemas de cómputo? A partir de las distintas caracterizaciones que se han propuesto (ver p. ej., Donald, 1991; Pinker, 1997; Rivière, 1991; Rivière y Núñez, 1996; Tomasello, 1999, 2000), tal vez muchas de esas diferencias cognitivas podrían resumirse no tanto en la capacidad de representarse el mundo, y actuar de acuerdo con esas representaciones, un rasgo que en realidad caracteriza a todos los sistemas cognitivos y los diferencia de otros sistemas meramente físicos y/o biológicos, sino en la capacidad *específica* (el sentido estricto del término ya que, hasta donde yo sé, es única) de *conocer nuestras propias representaciones* y de este modo poder actuar sobre ellas, potencialmente controlarlas e incluso modificarlas. Entendido así, como un sistema de conocimiento, dotado no sólo de representaciones, sino de representaciones de esas representaciones (o metarrepresentaciones) el sistema cognitivo humano, la mente de Hal, a diferencia del resto de los mecanismos cognitivos que conocemos (¿acaso ellos nos *conocen* a nosotros?), dispondría no sólo de *procesos y representaciones implícitos*, como los que tienen esos otros mecanismos, sino también

1. En realidad, a pesar de su extraordinario éxito el programa mecanicista en biología tampoco ha logrado otra de sus promesas esenciales: lograr explicar (o replicar) la vida (Lewontin, 2000). De hecho, no es casualidad que, al menos según Carims-Smith (1996), explicar la vida y la conciencia son los dos grandes problemas pendientes para la ciencia mecanicista.

de *procesos y representaciones explícitos*, aceptando la definición de Anderson (1996, pp. 123-124), según la cual son «procesos explícitos aquellos de los que se puede informar y procesos implícitos aquellos de los que no (se puede informar)». Para encontrar a Hal deberíamos ir por tanto más allá del enfoque cognitivo, y, sin necesidad de abandonar el nivel de análisis representacional, definir un nuevo objeto de estudio, o nivel de análisis, en términos de conocimiento. En lugar de reducir las representaciones a información (cómputos binarios entre unidades de información) debemos ampliar las representaciones en conocimiento, lo que requiere recuperar para la psicología cognitiva aquellos viejos problemas del contenido, la conciencia y la cultura en las representaciones.

La psicología del conocimiento: el contenido, la conciencia y la cultura en las representaciones

A pesar de su notable éxito, el paradigma cognitivo clásico del procesamiento de información ha entrado en crisis en la medida en que, a sus ya más de cuarenta años desde su nacimiento allá por 1956, aún es incapaz de dar cuenta de muchas de las funciones cognitivas que se le suponen a Hal (ver de Vega, 1998), entre las que destacan sin duda sus limitaciones en usar e intercambiar significados a través del lenguaje, emocionarse y sentir empatía, aprender de un modo no mecánico o repetitivo y, desde luego, disponer de intenciones y conciencia de sí mismo. Todas estas competencias cognitivas, y algunas otras *específicas* de la mente humana, y supuestamente también de la de Hal (el sentido del humor, la estética, el engaño, etc.), resultan en buena medida ajenas aún a ese sujeto computacional. Y lo que tienen en común es que requieren tener *conocimiento* y no sólo procesos cognitivos (percepción, atención, aprendizaje, memoria, etc.) eficaces. Según la terminología de Dienes y Perner (1999) el conocimiento implicaría tener una actitud proposicional con respecto a una representación, es decir, poder predicar algo con respecto a ella, lo que sin duda está fuera de las capacidades de todos los sistemas cognitivos artificiales (y que, sin embargo, recuérdese, era lo que caracterizaba *específica/mente* a Hal). Aunque con menos certeza, también parece fuera de las posibilidades cognitivas del resto de los sistemas cognitivos naturales (p. ej., Tomasello, 1999), por lo que sería una competencia cognitiva exclusiva de la mente humana (Pozo, 2001). Mientras que compartiríamos con el resto de los sistemas cognitivos la capacidad de generar representaciones implícitas mediante procesos igualmente implícitos, la cognición explícita, la capacidad de explicitar nuestras representaciones y, en menor medida, nuestros procesos (de Vega, 1995), sería una competencia cognitiva que abriría nuevos horizontes cognitivos a la mente humana. Curiosamente, aunque la psicología cognitiva parece haber descubierto muy recientemente la relevancia de la cognición implícita (p. ej., Froufe, 1996; Kirsner *et al.*, 1998; Reber, 1993; Rosetti y Revonsuo, 2000; Underwood, 1996), la práctica totalidad de la investigación cognitiva, sea clásica o conexionista, ha estado centrada en nuestra mente implícita, en la manipulación de representaciones, y es muy poco lo que

sabemos todavía sobre las funciones cognitivas necesarias para convertir esas representaciones en conocimiento, es decir sobre la cognición explícita. Sin embargo, comienza a haber indicios, en el marco de los ya clásicos estudios sobre disociaciones cognitivas, de que lo que *disocia* a la mente humana de otros sistemas cognitivos, ya sea en términos de memoria, aprendizaje, lenguaje, etc., es precisamente la capacidad de conocer, de explicitar sus representaciones (Dienes y Perner, 1999; Donald, 1991, 2001; Karmiloff-Smith, 1992; Pozo, 2001).

Según este argumento, el sujeto tradicional de la psicología cognitiva, el sujeto teórico pero sobre todo el sujeto experimental, ha estado privado, por principio, de conocimiento. Si se analizan la mayor parte de las tareas utilizadas en los laboratorios de la psicología cognitiva experimental se comprobará que se han diseñado de forma que se priva al sujeto de su conocimiento sobre la situación (mediante estímulos neutros, información sin significado, tareas descontextualizadas, etc.), por lo que se elimina en lo posible, como una posible perturbación, el conocimiento que el propio sujeto pudiera tener sobre sus propias representaciones y procesos. Ello tiene ventajas metodológicas que no debemos soslayar, ya que evitar el control explícito de la tarea por parte del sujeto permite un mayor control experimental de la situación (de hecho permite en buena medida separar el nivel de análisis cognitivo de la propia fenomenología), pero paga un precio que tampoco debemos olvidar. No creo que exagerara Oliver Sacks (1985) en su célebre relato *El hombre que confundió a su mujer con un sombrero*, cuando concluía que ese hombre (el doctor P), que sufría una forma sumamente extraña de *agnosia*, era una buena metáfora de la propia psicología cognitiva del procesamiento de información. El doctor P. extraía información del entorno pero era incapaz de integrar la información perceptiva que procesaba en un todo con significado. Percibía detalles, formas abstractas, pero no era capaz de situarlos en el mundo, de darles vida y color, en definitiva de conocer sus propias representaciones implícitas. Según Sacks (1985, p. 41 de la trad. cast.) «por una especie de analogía cómica y terrible, la psicología y la neurología cognitiva de hoy se parecen muchísimo al pobre doctor P. Necesitamos lo real y concreto tanto como lo necesitaba él; y no nos damos cuenta, lo mismo que él. Nuestras ciencias cognitivas padecen también una agnosia similar en el fondo a la del doctor P. El doctor P. puede pues servirnos de advertencia y parábola de lo que sucede a una ciencia que evita lo relacionado con el juicio, lo particular, lo personal y se hace exclusivamente abstracta y estadística».

¿Pero qué le faltaría al sujeto cognitivo, ya sea aquel de las computaciones sobre representaciones, o el más reciente y dinámico sujeto conexionista, para tener, como Hal, conocimiento? ¿Por qué padece esa agnosia teórica y experimental que denuncia Sacks? ¿Qué elementos debemos añadir a su dieta cognitiva para que además de representar o modelar el mundo lo sienta, lo padezca y en suma lo conozca? En mi opinión hay tres componentes esenciales, ya mencionados, que están ausentes de la dieta, sin duda *light*, a la que ha estado sometido el sujeto computacional, una dieta *sin* contenidos, *sin* conciencia y *sin* cultura (Pozo, 2001), tres componentes sin los cuales pueden producirse algunas formas implícitas de representación cognitiva, pero no es posible el conocimiento.

Comenzando por la primera de las carencias, la falta de contenidos, el origen computacional, y por tanto estadístico y formal, del sujeto cognitivo ha hecho que

el contenido, y en suma el significado, del procesamiento se subordinara casi siempre a la forma de esos cálculos. El problema es cómo puede surgir la semántica de unos cálculos exclusivamente sintácticos (Glenberg, 1997; Mateos, 1995; Pozo, 1989; Rivière, 1991; Russell, 1984). En suma, se trata del viejo problema del significado, que posiblemente supone el más serio problema que debe afrontar la psicología cognitiva (Gómez Crespo, 1998) y que posiblemente podría constituir el verdadero «salto evolutivo» entre otros sistemas cognitivos y la mente humana. La solución, tanto clásica como conexionista, ha sido intentar reducir el significado a una cuestión de cantidad de cálculo, dentro de la lógica de reducir las representaciones a información. El problema es que aumentos masivos en la cantidad de cálculo, aunque pueden mejorar considerablemente la eficacia cognitiva no traen consigo ni siquiera un atisbo de significado (o de conocimiento). Como el doctor P., el sujeto de la psicología cognitiva dispone de representaciones y procesos implícitos muy complejos y eficaces, pero totalmente carentes de contenido, de significado referencial. Lo que la psicología computacional, sea clásica o conexionista, debe explicar es cómo es posible que sistemas con capacidades de cálculo infinitamente superiores a la mente humana (como el célebre *Deep Blue II* que hace unos años venciera a Kasparov) no dan muestra alguna de tener conocimiento o de comprender lo que hacen, como haría una mente humana experta en esa misma situación, mucho menos dotada en cantidad de cálculo pero mucho más eficaz y selectiva en su procesamiento (de hecho, la verdadera cuestión cognitiva no es cómo pudo la máquina ganar a Kasparov sino cómo es posible que le resultara tan difícil ganarle, dadas las abismales diferencias de cálculo entre ambos).

Lo que diferencia, o disocia, a la mente humana, y a Hal, de esos otros sistemas es la naturaleza cualitativa y no la cantidad de esos cálculos. No es reduciendo las representaciones a información como vamos a poder abordar el problema del significado. Mientras que la información puede entenderse como el número de opciones generadas ante una serie de variables binarias, una representación implica el uso funcional de esa información codificada como sustituto de otro suceso (Denis, 1991). Para que haya una representación tiene que haber no sólo forma, reglas estadísticas, sino un contenido, un mundo al que esa representación se refiera (de Vega, 1995). Como mostrara la célebre parábola de la «habitación china» de Searle, el sujeto cognitivo clásico es un sujeto solipsista, encerrado en su propio mundo (podríamos decir representacional, pero en realidad informativo) que carece de referentes externos, de mundo al que adaptar sus representaciones. Como señala Russell (1984, p. 91) «los símbolos que manipulan los computadores no tienen significado en el sentido de referirse a alguna cosa del mundo externo sino que tienen significado en el sentido de tener un papel formal que desempeñar en el programa, un significado funcional». En este sentido, la metáfora computacional, por su naturaleza tecnológica y no biológica, ha oscurecido buena parte de las soluciones². Es difícil que el sujeto com-

2. En mi opinión, la solución *fodoriana* (el «cognitvista completo» según Gardner, 1985) al problema del significado es bastante significativa del callejón sin salida de la psicología computacional: explica las funciones cognitivas básicas (implícitas) en términos de módulos innatos, preprogramados genéticamente, pero sin proponer ninguna explicación que dé cuenta del origen y significado adaptativo de esos módulos mentales, desde un punto de vista evolucionista (Tomasello, 1999), al tiempo que niega la posibilidad de estudiar científicamente las funciones explícitas, y con ello se destierra una vez más el significado, el conocimiento, y por tanto a Hal, del paraíso cognitivo.

putacional comprenda lo que hace, que sus operaciones cognitivas tengan contenido, significado, si en realidad esas operaciones sólo tienen un contenido computacional (ver Russell, 1984).

La psicología cognitiva es y ha sido una psicología de los procesos (Tulving, 1989), que sólo muy secundariamente, y en ámbitos concretos como la psicología del pensamiento (Holland *et al.*, 1986; Nisbett, 1993; Pozo, 1988; Woll, 2002), o el aprendizaje y la instrucción (Pozo, 1989; Voss, Wiley y Carretero, 1995) ha aceptado que esos procesos se ejercen sobre ciertos contenidos representacionales, es decir sobre un mundo (social, físico, etc.) que no sólo restringe el funcionamiento de esos procesos sino que en buena medida debe explicarlos. Como señala Pinker (1997) la estrategia de la psicología cognitiva debe ser la de una «ingeniería al revés», ya que en lugar de tener que diseñar, como hacen los ingenieros, un sistema que cumpla ciertas funciones cognitivas (p. ej., moverse eficazmente en un entorno cambiante, jugar al ajedrez, usar referencialmente el lenguaje, aprender de modo comprensivo, etc.), disponemos ya de ese sistema, y al nivel más eficiente imaginable (la mente humana), pero no sabemos con certeza qué funciones cognitivas cumple, para qué sirve. Mientras los ingenieros tienen un problema y deben buscar una solución, la psicología tiene a su disposición un sistema que resuelve casi todos los problemas cognitivos relevantes (desplazarse y anticipar el movimiento de los objetos, integrar diferentes fuentes de información sensorial en una representación única, producir y comprender el lenguaje, atribuir intenciones y deseos a sí mismo y a los demás, dar significado a sus representaciones, reorganizar sus conocimientos, etc.), pero aún no hemos llegado a comprender cómo resuelve la mente humana muchos de esos problemas y, en muchos casos, ni siquiera sabemos cuáles son exactamente los problemas técnicos o teóricos que resuelve. Por ejemplo, la psicología cognitiva no tiene que imaginar un sistema cognitivo capaz de conocer, no tiene que diseñar a Hal, porque Hal surge de modo natural en cada uno de nosotros, pero ¿cómo y por qué tiene Hal conocimiento y no sólo representaciones? ¿qué ventajas adaptativas proporciona el conocimiento con respecto a las representaciones implícitas para que la mente humana haya evolucionado como un sistema de conocimiento?

Estas preguntas, necesarias si creemos realmente que el sistema cognitivo es un producto de la historia evolutiva, han estado clamorosamente ausentes de la psicología cognitiva (de Vega, 1998) y, en mi opinión, sólo pueden responderse devolviendo al sujeto cognitivo al mundo «real», analógico y tridimensional, físico pero sobre todo social, causal pero también intencional, en el que surgió la mente humana. Los «cómputos sobre representaciones» del sujeto cognitivo no pueden referirse únicamente a espacios artificiales de laboratorio, sino a ese mundo físico y social del que esa mente surgió. De entre las posibles formas de devolver al sujeto cognitivo al mundo, y en definitiva a la propia historia evolutiva de la mente humana, una hipótesis muy sugerente es la de la *embodied cognition*, la idea de que nuestras representaciones y procesos cognitivos están *encarnados*³ en un cuerpo que es a su vez producto de una historia evolutiva, y que la forma en que

3. Traduzco *embodied* por *encarnado*, ya que la traducción más precisa, *incorporado*, está ya descontada, o si se quiere «incorporada», al propio lenguaje cotidiano, con lo que ha perdido su fuerza denotativa. Por lo demás, la imagen de una mente encarnada trae para muchos de nosotros connotaciones muy personales.

ese cuerpo (que lleva una mente *incorporada*) procesa y se representa un mundo físico y social dado proporciona el contenido y significado esencial, primario, de nuestras representaciones (Clark, 1997; de Vega, 2002; Edelman, 1987; Glenberg, 1997; Pinker, 1997; Pozo, 2001). Ese significado primordial de las representaciones implícitas surgiría de la forma en que nuestro cuerpo, en el que está literalmente incorporada nuestra mente, interactúa con el mundo. Quizás quien mejor haya expresado esta idea sea Antonio Damasio en su propuesta de superar *El error de Descartes*, el dualismo que tan profundamente atraviesa también a la psicología cognitiva del *hardware* y el *software*⁴, dentro del cual el problema del contenido, y del significado, me temo, no tiene solución. Según Damasio (1994, p. 213 de la trad. cast., énfasis del autor): «*la naturaleza dio con una solución muy efectiva: representar el mundo externo en términos de las modificaciones que causa en el cuerpo propiamente dicho, es decir, representamos el ambiente mediante las modificaciones de las representaciones primordiales del cuerpo propiamente dicho siempre que tiene lugar una interacción entre el organismo y el ambiente*». De esta forma, no es sólo que para comprender el problema del significado de nuestras representaciones debamos situar la mente en un cuerpo y a su vez éste en el mundo, sino que más allá de eso, debemos comprender de qué formas está representado el cuerpo –y con él el mundo– en nuestra mente (Pinker, 1997; Pozo, 2001).

En esta línea, Glenberg (1997) ha propuesto un modelo de cognición encarnada con una naturaleza implícita y basada en representaciones analógicas que conservarían las propiedades topológicas del mundo, un mundo que estaría representado así en buena medida por las «pautas de acción que nuestro cuerpo» puede ejercer sobre él. Según este modelo, la alternativa a la falta de significado referencial de las representaciones simbólicas del cognitivismo clásico es asumir que «la memoria se desarrolló al servicio de la percepción y la acción en un entorno tridimensional, y que la memoria está encarnada para facilitar la interacción con el entorno» (Glenberg, 1997, p. 1). De esta forma, nuestra mente tendría literalmente *incorporados* ciertos contenidos, que corresponderían a la forma en que el mundo se proyecta en nuestras representaciones a través de las estructuras orgánicas, corporales, que son a su vez un producto de la evolución y por tanto han sido seleccionadas por ese mismo mundo (Pinker, 1997).

En mi opinión, no sería necesario suponer que esos contenidos, esas representaciones específicas de ciertos rasgos estructurales de nuestro mundo, están contenidos en nuestra «memoria genética» en forma de paquetes informativos listos para ser activados, como suponen los partidarios de la existencia de módulos específicos de conocimiento (p. ej., Hirschfeld y Gelman, 1994; Pinker, 1997; Spelke y Newport, 1997; véase para un análisis crítico Gómez Crespo y Núñez, 1998). Al contrario, la concepción de la mente encarnada nos permite evitar el siempre incómodo recurso a los conocimientos innatos, ya que sería la forma en que nuestro cuerpo (de cuya conformación genética básica nadie puede dudar) interactúa con el mundo la que restringiría los contenidos esenciales de nuestras representaciones en ciertos dominios específicos. ¿Y cuáles son esos

4. Y que en mi opinión, a pesar de apoyarse supuestamente en la «metáfora del cerebro» tampoco resuelve el nuevo enfoque conexionista, tal como hasta ahora se ha definido (Pozo, 2001).

dominios? Desde una perspectiva evolucionista se ha argumentado que la mente humana, lejos de estar vacía de contenidos, *contiene* dispositivos cognitivos específicos para representar determinados dominios del mundo. Podríamos decir que esos dispositivos, sobre cuya naturaleza específica aún no sabemos lo suficiente, constituirían nuestro «equipamiento cognitivo de serie» (Pozo y Gómez Crespo, 2002) o estándar (Pinker, 1997) en esos mismos dominios. Aunque, desde luego, no hay acuerdo sobre la naturaleza y el número de esos dominios (ver p. ej., Hirschfeld y Gelman, 1994; Karmiloff-Smith, 1992; Sperber, Premack y Premack, 1995; Wellman y Gelman, 1997), al menos suelen reconocerse dos dominios básicos diferentes, el mundo físico y el mundo social, para los que la mente humana probablemente disponga de un «equipamiento cognitivo de serie» que contenga representaciones especializadas, que surjan de modo necesario de las interacciones encarnadas con ese mundo, un mundo que de algún modo seleccionó esa mente en un largo proceso evolutivo (o para ser exactos, como nos recuerda Pinker, 1997, fue la mente de nuestros antepasados, que aún habita en gran medida en nosotros como un vestigio arqueológico, la que fue seleccionada por un mundo físico y social que ya sólo existe en parte, ya que ha sido radicalmente alterado por la cultura).

La forma en que nuestro cuerpo nos informa de los cambios que se producen en el mundo físico restringe y da contenido a la representación de ese mundo material en nuestra mente. Vivimos en un mundo dotado de gravedad y de múltiples fricciones imperceptibles que han restringido nuestra propia estructura física y a través de ella los contenidos primordiales de nuestra física intuitiva, es decir nuestras representaciones implícitas sobre la materia y sus cambios (Pozo y Gómez Crespo, 1998, 2002). Nuestra física intuitiva estaría organizada en forma de teorías implícitas contenidas en nuestra memoria, basadas en ciertos *principios* o restricciones internas, algunas de las cuales estarían activas ya desde la primera infancia (p. ej., Spelke, Phillips y Woodward, 1995). Esta física intuitiva, de naturaleza implícita y encarnada en nuestra mente, estaría muy alejada de la física científica y sería muy difícil de modificar, incluso mediante la instrucción (Gómez Crespo y Pozo, 2001; Pozo y Gómez Crespo, 1998; Pozo, Gómez Crespo y Sanz, 1999).

Pero hay otras representaciones primordiales o encarnadas, además de las que se refieren al mundo físico, que también tienen su origen en nuestro cuerpo. La vida en grupo, característica de los primates, debió de hacer necesario desde muy pronto no sólo representar los cambios que el ambiente físico producía en el cuerpo, sino también los cambios que en ese cuerpo producían las acciones de otros congéneres. Damasio (1994) habla de «marcadores somáticos» para referirse a las respuestas emocionales, muchas veces implícitas o encubiertas, que acompañan a las consecuencias de nuestras acciones. Es bastante probable que la selección natural nos haya preparado no sólo para percibir esos marcadores somáticos internos en nuestro propio cuerpo, sino también para *simularlos* o detectarlos en los de los demás (Humphrey, 1992; Pinker, 1997; Tomasello, 1999). Una mente *que contenga* dispositivos para generar representaciones de esos marcadores somáticos ajenos y anticipe sus consecuencias tendría un alto valor adaptativo. De hecho, así parece aceptarse hoy en día: desde muy temprano en su evolución

cognitiva, los seres humanos tuvieron que volverse psicólogos naturales, especialmente dotados para procesar la conducta de los demás en términos de representaciones mentalistas. Las personas tendríamos, además de esa física intuitiva, también una psicología intuitiva, que se constituiría en una teoría implícita con un contenido específico, lo que se ha dado en llamar *teoría de la mente*.

El contenido implícito de esa psicología intuitiva, si nos atenemos a los datos de las investigaciones recientes, en cualquiera de sus variantes, y sin entrar a debatir entre ellas (p. ej., Dienes y Perner, 1999; Flavell, 1999; Gómez Crespo y Núñez, 1998; Leslie, 1987; Martí, 1997; Rivière, 1991; Rivière y Núñez, 1996; Tomasello, 1999), sería precisamente aquella capacidad de generar un tipo muy específico de representaciones, o de *metarrepresentaciones*, que implican adoptar *actitudes proposicionales* con respecto a nuestras representaciones y las de los demás, es decir convertir nuestras representaciones en objeto de representación, en conocimiento. Hay muchos sistemas biológicos e incluso mecánicos capaces de almacenar información, incluso muchos sistemas cognitivos capaces de representar o modelar mundos concretos, pero sólo las personas tenemos *conocimiento*, porque sólo las personas recordamos, creemos, anhelamos, sabemos, dudamos, ignoramos, sólo las personas adoptamos *actitudes proposicionales* con respecto a las representaciones del mundo que tenemos en nuestra mente. Esa capacidad de convertir las propias representaciones (implícitas) en objeto de representación (explícita), esencial para elaborar una psicología del conocimiento, estaría posiblemente vinculada al desarrollo de las capacidades mentalistas propias de nuestra especie (Gómez Crespo y Núñez, 1998; Humphrey, 1992; Rivière, 1991), que no sólo nos permite sino que nos induce a atribuir esas mismas competencias mentalistas (intenciones, conciencia, emociones, etc.) a los demás, y por tanto, según la argumentación aquí desarrollada, a atribuirles conocimiento, actitudes proposicionales con respecto a sus representaciones.

Esa psicología intuitiva, con sus competencias mentalistas, es para muchos autores (p. ej., Gómez Crespo y Núñez, 1998; Humphrey, 1992; Tomasello, 1999) el mecanismo evolutivo que permitió desarrollar en la mente humana esa capacidad de conocer o explicitar las representaciones, y no sólo uno de los productos del desarrollo de esa misma capacidad. Esta hipótesis resulta creíble, pero aunque no fuera así, la diferencia cualitativa esencial entre la cognición humana y la de otros sistemas cognitivos parece en todo caso claramente vinculada a la especificidad de nuestras representaciones y procesos cognitivos explícitos, que hacen posible el conocimiento (Pozo, 2001). Las investigaciones recientes sobre la disociación entre el funcionamiento implícito y explícito de nuestra mente, están revelando el papel esencial de los procesos explícitos en el desarrollo pleno de nuestras funciones cognitivas (p. ej., Dienes y Perner, 1999; Kirsner *et al.*, 1998; Rosetti y Revonsuo, 2000; Underwood, 1996). La primacía filogenética y ontogenética de los procesos implícitos sobre los explícitos, postulada por Reber (1993) pero también desde perspectivas distintas por Donald (1991) o Karmiloff-Smith (1992), es compatible con un modelo de integración jerárquica en el que las primeras funciones (implícitas) son condición necesaria para el desarrollo de los sistemas posteriores (explícitos), pero éstos a su vez reconstruyen el funcionamiento cognitivo de esos niveles más primarios (Pozo, 2001; Pozo y Gómez Crespo, 1998).

De esta forma, el estudio de los procesos implicados en la generación y adquisición de conocimientos en dominios específicos debe apoyarse en los procesos cognitivos que dan lugar a las representaciones implícitas, encarnadas, en esos mismos dominios, parte de los cuales han sido investigados por la psicología cognitiva en estas últimas décadas. Aunque la psicología cognitiva haya descubierto hace bien poco, y no sin sorpresa, los fenómenos disociativos, y con ellos la cognición implícita, la mayor parte de las investigaciones cognitivas han estado de hecho centradas en esas funciones no conscientes. Como señalan Neal y Hesketh (1997, p. 34), la paradoja es que «prácticamente todos los modelos de procesamiento de información parecen asumir que las formas complejas de procesamiento de información ocurren fuera de la conciencia, pero por otra parte muchos psicólogos cognitivos se han mostrado muy escépticos ante fenómenos tales como el aprendizaje implícito». Tal vez la paradoja no sea tal si entendemos que añadir el adjetivo «implícito» a procesos tan aceptados ya como la memoria, el aprendizaje, la percepción o incluso el razonamiento (p. ej., Kirsner *et al.*, 1998; Rosetti y Revonsuo, 2000) supone (aunque sea implícitamente, otra paradoja) aceptar también una versión explícita de esos mismos procesos, cuya función y naturaleza cognitiva, para ser relevante la distinción, debería ser cualitativamente distinta a la de sus equivalentes implícitos. Lo cierto es que aún hoy es muy poco lo que la psicología cognitiva puede decirnos sobre el funcionamiento consciente o explícito de cada uno de esos procesos (de Vega, 1995). Incluso quienes defienden la existencia de procesos cognitivos implícitos (p. ej., Kirsner *et al.*, 1998; Reber, 1993) carecen, en general, de una teoría sobre la función cognitiva de los procesos explícitos, que parecen limitarse, en muchos casos, a encender la luz en la habitación oscura del conocimiento (Pozo, 2001).

Sin embargo, si seguimos situándonos en una perspectiva evolucionista (la única posible no sólo para la psicología del conocimiento sino para la propia psicología cognitiva) los procesos de explicitación o conciencia deben tener alguna función cognitiva y por tanto no pueden ser, como parece haberse asumido (implícitamente, de nuevo) un epifenómeno, causalmente irrelevante (de Vega, 1998). Para la psicología cognitiva clásica, la conciencia o los procesos explícitos han sido, por recordar la metáfora usada en su momento por William James (1890), la espuma de los procesos cognitivos. Ésta es una idea que, de hecho, la psicología científica ha heredado de la propia tradición de la ciencia mecanicista en la que se inscribe. Hace más de un siglo, el propio Thomas Huxley, quizás el primer discípulo de Darwin, ya sostenía que la conciencia «es simplemente un producto colateral, tan carente de poder como el silbido de vapor que acompaña al funcionamiento de la locomotora y que no tiene ninguna influencia en su maquinaria» (citado por Humphrey, 1986, p. 62 de la trad. cast.). Según esta idea, las funciones conscientes son *sólo* como un efecto de los mecanismos mentales (en el caso de la psicología cognitiva al uso, como un mero efecto de los cómputos realizados por la mente) pero no como una posible causa, a su vez, del funcionamiento de esos mismos mecanismos. Pero si la materia llegó a organizarse de forma tal que produjo esos efectos subjetivos de la conciencia, éstos deben tener alguna función en la propia organización de esa materia, deben de tener, por decirlo con la contundencia de Cairns-Smith (1996), un efecto físico sobre ella. O, como dice Lewontin (2000), la célebre invitación de Lau-

ren Bacall a Humphrey Bogart en *Tener y no tener* («si me necesitas, silba») no serviría de mucho en esa «locomotora computacional», la máquina de vapor cognitiva imaginada por el procesamiento de información.

Tal vez todos coincidiremos en que una de las carencias esenciales del sujeto cognitivo para emular a Hal sea precisamente la endeblesz de sus procesos conscientes, y más concretamente aquellas formas de conciencia que tienen que ver con la capacidad de representar las propias representaciones, la función cognitiva de metarrepresentación (Rivière, 1997; Sperber, 2000). Dado el carácter polifacético de la conciencia, que se ha desvelado no como un proceso psicológico, sino como una familia de procesos, no siempre bien avenida (p. ej., Damasio, 1999; de Vega, 1995; Donald, 2001), existen sin duda otras formas de conciencia que sí son accesibles al sujeto cognitivo (en forma de memoria de trabajo, asignación y control de recursos cognitivos, etc.), pero uno de los tipos de conciencia que le ha sido negado a ese sujeto cognitivo, debido seguramente a su sabor fenomenológico, tan poco grato a la psicología experimental, es el acceso consciente a sus propias representaciones, es decir la capacidad de convertir sus representaciones implícitas en representaciones explícitas o conocimiento propiamente dicho, asumiendo al mismo tiempo que esa explicitación es un proceso transformador o constructivo, que modifica la naturaleza de esas representaciones explicitadas. En el marco de una psicología del conocimiento, la función de los procesos explícitos sería posiblemente la de favorecer la *redescripción representacional* de aquellos productos de la cognición implícita (Karmiloff-Smith, 1992), de forma que no sólo sean más accesibles al propio sistema cognitivo sino también más fáciles de relacionar con otras representaciones, generando nuevas formas de conocimiento. En este sentido, la conversión de las representaciones en conocimiento mediante su explicitación progresiva reduciría la dependencia del sujeto con respecto al ambiente, ya que le permitiría nuevas formas de organizar sus representaciones, o su conocimiento, más allá de la estructura que los estímulos adoptaran en el mundo (Donald, 1991). Los sistemas que tienen conocimiento, como Hal, tienen por tanto una mayor autonomía funcional en la gestión de sus propias representaciones y en suma en su interacción con el mundo, lo que les hace capaces de generar y adquirir nuevas formas de conocimiento que se alejan cada vez más de la apariencia inmediata (o mejor mediada por las representaciones encarnadas) del mundo, lo que permite no sólo una mayor flexibilidad cognitiva y, en consecuencia nuevas adaptaciones al mundo físico y social representado en la mente, sino incluso convertir algunos de esos mundos posibles, imaginados o representados, en mundos reales, contruidos a través del conocimiento así generado.

Ahora bien, para que haya conocimiento y este conocimiento nos permita generar nuevos mundos posibles, mentales, en los que vivir no basta con esas funciones conscientes. La explicitación requiere no sólo conciencia sino también un código o lenguaje externo a la propia representación al que traducir ésta para poder compartirla y distribuirla socialmente. Según ha argumentado Tomasello (1999; también Tomasello, Kruger y Rafter, 1993) un rasgo diferencial del conocimiento humano con respecto al de otras especies (si es que lo tienen, que ya es en sí dudoso) es su carácter acumulativo. El conocimiento, con todas sus consecuencias cognitivas para la mente humana, sólo puede entenderse en el marco de

una transmisión, de generación en generación, de las representaciones que la cultura ha ido no sólo explicitando sino conservando gracias al desarrollo (cultural pero también cognitivo) de sistemas de representación o memoria externa. Aunque obviamente podemos asumir otros análisis sobre las relaciones entre mente y cultura, desde la perspectiva cognitiva, o representacional, que aquí estamos adoptando, podemos aceptar, con Sperber (1996) que la cultura es el conjunto de «representaciones asociadas y ampliamente distribuidas» entre una población. Así, analizar la influencia de la cultura en la mente es estudiar cómo esas representaciones públicas o compartidas se convierten en representaciones mentales (y también, como veremos, a la inversa, cómo el funcionamiento cognitivo restringe o formatea el desarrollo de esas representaciones culturales). A diferencia de lo que sucede con otras especies, la cultura humana ha permitido una adaptación cognitiva lamarckiana en vez de darwiniana, de forma que los caracteres (culturalmente) adquiridos sí se transmiten (culturalmente) a las nuevas generaciones (Pinker, 1997). O, dicho en palabras del paleontólogo Juan Luis Arsuaga (2001, p. 356), «hoy por hoy, el único caso en el que estamos seguros de que el comportamiento no programado genéticamente cambió las condiciones del ambiente, es decir, modificó el nicho ecológico primero y más tarde el propio medio físico, creando presiones nuevas, es el de la tecnología humana». De esta forma debemos pensar las relaciones entre mente y cultura en términos de una «co-evolución».

Los conocimientos que cada uno de nosotros adquirimos reformatean de algún modo nuestra mente haciéndola sin duda distinta a la mente no sólo de *nuestros primos* los primates sino incluso de nuestros propios antepasados. Las formas de aprendizaje explícito, apoyadas en la cultura y en la instrucción, nos diferencian así del funcionamiento de otros sistemas cognitivos, haciendo posible no sólo el conocimiento, sino también su adquisición y transmisión. El conocimiento no sólo requiere hacer explícitas las representaciones, sino que a su vez esa explicitación se apoya en sistemas externos de representación que hacen posible transmitir y adquirir esos conocimientos. Como ha mostrado Donald (1991), la construcción de la propia mente humana, en sus orígenes, implica de algún modo un proceso de interiorización, o mentalización, de esos sistemas culturales de representación que dan lugar a diferentes mentes o sistemas cognitivos (episódico, mimético, simbólico y teórico). No será posible por tanto entender los procesos de explicitación y adquisición del conocimiento sin situarlos en ese marco de transmisión cultural, que en nuestra sociedad se basa en sistemas de instrucción más o menos formal.

Con respecto a las posibles relaciones genéticas mutuas entre esas funciones conscientes y los sistemas de representación explícita en que se apoyan, en especial el lenguaje, aunque hay posiciones muy diversas (p. ej., Arsuaga, 2001; de Vega, 1995; Karmiloff-Smith, 1992; Pinker, 1997; Rivièrè, 1991), tal vez podamos asumir (p. ej. con Donald, 1991, o Tomasello, 1999) que las funciones conscientes, en una forma primitiva, sirvieron como preadaptaciones evolutivas o *exaptaciones*, oportunamente colonizadas por ese producto cultural tan complejo que es el lenguaje. Hay incluso quien cree, como Gould (1991), que ambos, tanto la conciencia como el lenguaje, son efectos colaterales, no seleccionados, del propio desarrollo del cerebro, aunque esta hipótesis parece aún menos probable. En todo

caso, si bien el lenguaje difícilmente puede ser el origen de la metarrepresentación ya que no parece creíble que el lenguaje pueda existir sin una función cognitiva explícita y/o una intención comunicativa previa (Tomasello, 1999), supone sin duda un multiplicador sumamente potente de esas capacidades metarrepresentacionales y de los conocimientos que a través de ellas se generan. Las representaciones explícitas, el conocimiento en suma, para ser efectivo, requiere de mediadores culturales, sistemas externos de representación dados por la cultura, que no sólo sirvan de soporte a esas nuevas formas de conocimiento, sino que sobre todo les proporcionen una nueva forma, las reformateen a la vez que las redesciben (Martí y Pozo, 2000). Como señala Bruner (1993) no se sabe si es la mente la que «adquiere» la cultura o viceversa, pero lo cierto es que esas representaciones culturales generan nuevas posibilidades representacionales y cognitivas. Y sin embargo, de la misma forma que la psicología cognitiva se ha desarrollado ajena a los contenidos y a las funciones conscientes, tampoco ha considerado, ni siquiera mínimamente, la función esencial de las representaciones culturales en la construcción de la mente humana. Si el sujeto de la psicología cognitiva ha sido experimentalmente agnóstico también ha sido experimentalmente inculto, es decir un sujeto privado de cultura (Cole, 1996). Estas dos carencias no son sino dos reflejos diferentes, pero convergentes, de la ausencia del *conocimiento* en la psicología cognitiva del procesamiento de información.

Una forma de situar no sólo la mente en la cultura, sino más allá de ello, como reclama Bruner (1993), la cultura en la mente, en el marco de una psicología del conocimiento, sería investigar de qué modo los sistemas de representación cultural median en esa redesccripción representacional de los contenidos de la mente implícita, ayudándonos a ir más allá de ese equipamiento cognitivo de serie, al que antes me referí, y que ha constituido el objeto esencial de la investigación cognitiva. De hecho, hay motivos para sospechar que algunos de los procesos y representaciones que ha estudiado tradicionalmente la psicología cognitiva —usualmente con una muestra tan homogénea y sesgada como los estudiantes de psicología— tienen en sí mismos un fuerte componente cultural, son de hecho *extras* o añadidos culturales (Pozo, 2001). Un ejemplo de ello serían, en opinión de W. Kintsch (1998), las propias formas de representación simbólica que han constituido el núcleo de buena parte de la investigación cognitiva clásica de «representaciones sobre computaciones». Según nos recuerda Kintsch (1998, p. 29), «no se puede tener pensamiento simbólico en una mente individual, sólo en una mente que es parte de una cierta cultura. Pensamos como pensamos porque generaciones anteriores han inventado los modos de pensar que ahora usamos y han acumulado mucho de lo que constituyen los contenidos específicos de nuestro pensamiento». Lo que no ha impedido que la psicología cognitiva haya asumido, sin siquiera abrir la puerta al debate, que las representaciones simbólicas pueden estudiarse sin referencia alguna a la cultura. Y como muestra un botón: en el notable manual de Baddeley (1997), *Memoria humana*, tal vez la mejor ventana para acercarse a la investigación reciente en este área, una de las más importantes en el desarrollo de la psicología cognitiva, la cultura no aparece no ya como nivel de análisis teórico, sino tampoco como capítulo o tema relevante. De hecho, ni siquiera aparece en el índice temático como des-

criptor o palabra-clave. La memoria que ha estudiado la psicología cognitiva es por tanto una *memoria sin cultura*, que es lo mismo que decir sin conocimiento. Otro tanto podríamos decir con seguridad del resto de los procesos cognitivos.

Sin embargo, los estudios que se han realizado sobre los *cambios* cognitivos producidos en la mente humana como consecuencia de la adquisición de diferentes sistemas culturales de conocimiento avalan la idea de que esa enculturación tiene unos efectos mucho más profundos sobre el sistema cognitivo de lo que la psicología computacional ha admitido. Así por ejemplo, en su brillante análisis sobre los efectos cognitivos de la alfabetización, Olson (1994) ha mostrado no sólo la estrecha dependencia entre las funciones cognitivas y los sistemas de representación cultural, sino cómo varias de las formas de representación mental estudiadas por los psicólogos cognitivos han sido en realidad generadas culturalmente. De hecho, cabe pensar que, por su propia naturaleza, los sistemas culturales de representación son el medio esencial para la redescritión representacional o conversión de las representaciones en conocimiento en la mente humana (Martí y Pozo, 2000). Si la mente humana dispone de una representación encarnada de algunas de las dimensiones básicas del mundo en el que vive (tiempo, espacio, número, causalidad, relaciones interpersonales, etc.), los sistemas culturales de representación y conocimiento en esos mismos ámbitos generan nuevas competencias cognitivas que se incorporan a los procesos cognitivos básicos al tiempo que reconstruyen culturalmente la mente humana (ver Pozo, 2001). Sólo una investigación que combine las distintas dimensiones de cambio en la mente humana (filogénesis, sociogénesis y ontogénesis) nos permitirá comprender los efectos de diversos sistemas culturales de conocimiento sobre el funcionamiento cognitivo (p. ej., Valsiner, 1992).

En suma, si queremos no ya comprender a Hal sino siquiera comenzar a estudiar su funcionamiento mental, debemos superar algunas de las restricciones teóricas y metodológicas impuestas por los modelos computacionales, tanto clásicos como conexionistas, que aquí se han analizado en términos de tres carencias esenciales en la dieta del sujeto teórico y experimental de la psicología cognitiva, que vive en un mundo artificial, *sin* contenidos, conciencia ni cultura y por tanto *sin* conocimiento. Si queremos conocer a Hal, debemos permitirle a él también conocer, incorporando estos componentes a la dieta teórica y experimental de la psicología cognitiva. Pero ello, además de abrir nuevas posibilidades, traerá consigo nuevos problemas teóricos y experimentales, algunos de los cuales se avanzan en el próximo apartado, que sirve de cierre a este artículo.

Estudiando experimentalmente a Hal: una agenda de problemas para la psicología del conocimiento

Según hemos visto, recuperar a Hal para la psicología cognitiva requiere convertir al sujeto cognitivo en sujeto de conocimiento, lo que implica cómo mínimo incorporar (una vez más, en sentido literal) a la investigación cognitiva los

contenidos o significados, la conciencia o explicitación de las representaciones y la cultura en forma de sistemas externos de representación. Pero incorporar al estudio de las representaciones estos componentes requiere dar cuenta de la diferencia cualitativa entre el funcionamiento cognitivo explícito e implícito. Y ello exige proponer marcos teóricos que permitan integrar los mecanismos primarios del procesamiento implícito, aquel equipamiento cognitivo de serie, mostrando cómo esos procesos implícitos restringen el funcionamiento cognitivo explícito, al tiempo que los procesos explícitos reconstruyen o gestionan a su vez esas representaciones implícitas o encarnadas. En mi opinión estas nuevas exigencias no requieren abandonar el enfoque cognitivo, al menos entendido en el sentido amplio que se mencionaba al comienzo, que incluiría no sólo el enfoque computacional sino otros enfoques cognitivos alternativos, más abiertos al estudio del conocimiento, pero sí resultan inviables en el estrecho marco de una psicología *exclusivamente* computacional. Estudiar a Hal requiere adoptar un enfoque de conocimiento y no sólo cognitivo, ya que estudiar las representaciones explícitas (o conocimiento) como representaciones con contenidos, conciencia y cultura plantea nuevas exigencias tanto teóricas como metodológicas a las que debemos enfrentarnos.

Así, estudiar los contenidos de la mente y no sólo sus procesos (o para ser exactos estudiar la forma en que la mente procesa determinados contenidos preferentemente sobre otros) requiere definir cuáles son los dominios de conocimiento que deben estudiarse. Sin embargo, los diversos autores no se ponen de acuerdo no ya con respecto a cuáles son los principales dominios psicológicos del conocimiento, sino ni siquiera con respecto a los criterios para definirlos. Ésta es una primera tarea pendiente para la investigación, establecer los principales dominios del conocimiento humano. Tal vez parte del problema en la delimitación de esos dominios sea que diferentes autores (p. ej., Gopnik y Metzoff, 1997; Hirschfeld y Gelman, 1994; Karmiloff-Smith, 1992; Sperber, Premack y Premack, 1995; Wellman y Gelman, 1997) se apoyan para hacerlo en criterios diferentes (epistemológicos, psicológicos, académicos o instruccionales). Y cabe pensar que los dominios de conocimiento cambien en su organización cognitiva a medida que la mente humana incorpora distintos sistemas culturales de conocimiento (las producciones lingüísticas, artísticas, científicas, matemáticas, religiosas, etc.). Por tanto la investigación de esos dominios de conocimiento requerirá adoptar un enfoque dinámico centrado en el estudio del *cambio*, en lugar del enfoque estático propio de la psicología cognitiva clásica. Serán necesarias investigaciones comparadas sobre los cambios filogenéticos en la representación implícita de diferentes dominios (p. ej., Tomasello, 1999), junto con investigaciones sobre el cambio ontogenético o evolutivo en esos mismos dominios (Wellman y Gelman, 1997), así como análisis culturales de las diferentes formas en que los sistemas externos de representación pueden ayudar a reformatear el conocimiento en esos dominios y con ello los propios dominios de conocimiento, sobre los que la investigación cognitiva es aún más escasa, aunque esté recibiendo una atención creciente (p. ej., Cole, 1996; Nisbett *et al.*, 2001; Sperber, 1996).

La propia idea de que los dominios cognitivos pueden verse afectados por la ontogénesis y la cultura supone aceptar, en contra de lo que asumen (¿otra vez

implícitamente?) la mayor parte de los modelos cognitivos, y desde luego de los supuestos del procesamiento de información (Reber, 1993), que los propios procesos cognitivos (percepción, atención, memoria, aprendizaje, etc.), y no sólo las representaciones, pueden ser modificados por la experiencia, por la cultura y en definitiva por el aprendizaje. Un problema adicional para aceptar esta naturaleza dinámica o flexible del sistema cognitivo en diferentes dominios es que muchas veces éstos, en ausencia de argumentos más convincentes, se justifican en la existencia de supuestos módulos cognitivos específicos para cada uno de esos dominios, aunque eso sí casi nunca apoyados en una investigación comparada que ayude a comprender el origen genético y la función adaptativa de esos módulos (Tomasello, 1999). El estudio de la naturaleza encarnada de las representaciones implícitas puede proporcionar una alternativa teórica a estas concepciones modulares compatible con esa flexibilidad cognitiva, de representaciones pero también de procesos, sin la cual Hal no es fácilmente comprensible (Pozo, 2001). En este sentido, más que hablar de cambios conceptuales o de contenidos en el conocimiento de dominio, podemos pensar en los cambios representacionales que tienen lugar en esos mismos dominios (Pozo y Rodrigo, 2001), es decir en los cambios que la explicitación produce en la naturaleza de las representaciones contenidas en nuestra memoria (Karmiloff-Smith, 1992; Pozo, 2001; Rodrigo y Correa, 2001), debidas en buena medida a la influencia cultural mediada por la instrucción.

El estudio de los contenidos y dominios de conocimiento requiere por tanto diferenciar entre representaciones implícitas y explícitas en esos distintos dominios, lo que puede ayudar a esclarecer la función cognitiva de la explicitación en la construcción de esos conocimientos. En este sentido, una seria dificultad metodológica es la diferenciación entre representaciones implícitas y explícitas. Aunque es fácil encontrar ejemplos extremos, en los que suelen apoyarse los clásicos trabajos sobre disociaciones cognitivas, ya sean como consecuencia de trastornos en ciertas funciones cognitivas, o bien disociaciones evolutivas o inducidas experimentalmente, esta visión dicotómica, que separa las funciones explícitas e implícitas, no debe oscurecer las continuas relaciones que deben existir entre ambas formas de representación, de forma que la dimensión implícito-explicito debería ser entendida más como un continuo con niveles intermedios o progresivos de explicitación que como una dicotomía (Pozo, 2001). Esos diferentes grados de explicitación dependerían de la existencia de diversos componentes representacionales que pueden ser explícitos o no (Dienes y Perner, 1999), de diversos niveles de redescrición o explicitación de las representaciones (Karmiloff-Smith, 1992) o de diferentes relaciones o estructuras conceptuales entre esas mismas representaciones que podrían ser jerárquicamente explicitadas (Pozo y Gómez Crespo, 1998). Por tanto, más que ante una dicotomía, cuyas partes deban ser disociadas, nos encontraríamos ante una variedad de relaciones entre las formas implícitas y explícitas de representación que, más allá de esos efectos disociativos clásicos, para ser realmente comprendidas deberían ser asociadas o integradas en un mismo marco teórico (Kirsner *et al.*, 1998; Pozo, 2001; Rosetti y Revonsuo, 2000). Si queremos comprender a Hal debemos por supuesto separar experimentalmente sus diferentes funciones cog-

nitivas (implícitas y explícitas) pero también integrarlas o relacionarlas experimentalmente, ya que en la mayor parte de los contextos cognitivos para las que esos procesos son funcionales el sujeto de conocimiento debe actuar de una forma más bien integrada.

Una forma de interpretar esta convivencia, en un mismo sujeto de conocimiento, de las representaciones implícitas y explícitas es comprender su diferente función cognitiva. En este sentido, resulta relevante la distinción de Kirsh y Maglio (1994) entre la función pragmática de ciertas representaciones (en este caso las implícitas), que se constituirían en esquemas de acción, y la función epistémica de otras representaciones (las explícitas) que servirían no para actuar sobre el mundo sino para conocerlo, en suma para actuar sobre las propias representaciones que el sujeto tiene del mundo. Esta orientación interna de la función epistémica de las representaciones explícitas, por contraposición a la orientación externa de las representaciones implícitas (Pozo, 2001), tiene sin duda un notable regusto vygotskiano, lo que puede ayudarnos a encontrar vías de conexión teórica entre la psicología cognitiva y los enfoques culturales, en el marco de una psicología del conocimiento. De hecho, una estrategia para estudiar los efectos de diferentes formas de explicitación o conocimiento sobre el funcionamiento cognitivo implícito puede ser estudiar esas relaciones en contextos culturales o instruccionales diferentes, que planteen demandas (explícitas e implícitas) diferenciadas. Dado que la instrucción es en nuestra sociedad el principal vehículo para las transformaciones culturales de la mente, estudiar los efectos de la instrucción sobre el conocimiento y, en general, sobre las funciones cognitivas ha sido una estrategia esencial en la reciente investigación en psicología del conocimiento (p. ej., Carretero *et al.*, 1991; Reigeluth, 1999; Voss, Wiley y Carretero, 1995). Obviamente, desde esta perspectiva, considerando los criterios para la elección de las muestras exigidos por la psicología del conocimiento, la investigación debería apoyarse preferentemente en diseños cuasiexperimentales, ya que esos cambios instruccionales en la mente requieren condiciones y tiempos largos que no pueden ser controlados ni simulados experimentalmente, lo que probablemente exige acercamientos metodológicos no reduccionistas, pero tampoco eclécticos, que permitan integrar diferentes niveles de análisis o perspectivas (Pozo y Rodrigo, 2001).

Aceptando la sugerencia de Saalhouse (1991), según la cual ser experto en cualquier dominio de conocimiento implica superar, o si se quiere suspender representacionalmente (Pozo, 2001), en el sentido de Leslie (1987) o Rivière (1997), algunas de las restricciones naturales impuestas por lo que aquí he llamado el «equipamiento cognitivo de serie», el estudio de los cambios cognitivos que se producen al adquirir determinados conocimientos específicos de dominio, y la forma en que esos cambios pueden generalizarse a otros dominios, puede ayudar a comprender mejor la forma en que esos conocimientos culturales ayudan a reconstruir la mente humana (Pozo, 2001). En todo caso, esa reconstrucción cultural de la mente, la conversión de esas formas de conocimiento cultural en procesos y representaciones cognitivas, en artefactos cognitivos, en el sentido de Cole (1996), estaría también restringida por la dinámica del propio sistema cognitivo implícito, por las representaciones encarnadas profundamente arraiga-

das en nosotros. La reconstrucción de estas representaciones primitivas a partir de la adquisición de esos sistemas culturales de conocimiento permitiría generar nuevos tipos de representación (Pozo y Rodrigo, 2001), pero en todo caso esa reconstrucción estaría restringida por el propio sistema cognitivo. Por usar la metáfora de Jackendoff, se trataría de construir nuevas representaciones con piezas del *lego*, no con plastilina: los sistemas de conocimiento generados por la mente humana tienen una notable versatilidad (son muchas las construcciones distintas que pueden hacerse con las piezas de un *lego*), pero tienen una cierta forma inicial (un cierto contenido primordial, encarnado), por lo que no son infinitamente maleables, no son formas vacías, de modo que, en contra de lo que ocurre en los sistemas conexionistas, vacíos de contenidos, ciertas construcciones son más probables que otras (Pozo, 2001). Por tanto, la redescipción representacional estará *mediada* no sólo por los sistemas culturales de conocimiento *a través* de los que se reconstruyen los procesos y representaciones mentales sino también en buena medida por la propia mente encarnada desde la que se realiza esa redescipción.

En definitiva, la psicología del conocimiento no debería estudiar sólo cómo la cultura (o la instrucción) reconstruye la mente, sino también a la inversa cómo la mente reconstruye la cultura en el proceso de adquisición de esos conocimientos o saberes culturales, de qué modo la mente humana restringe el propio desarrollo cultural del conocimiento. Pero mientras la investigación instruccional ha sido abundante en estos últimos años, existen muchos menos estudios sobre los efectos de diferentes tipos de enculturación sobre el funcionamiento cognitivo en dominios específicos de conocimiento, así como también sobre la propia historia cultural de esos sistemas de conocimiento y la forma en que han evolucionado en función de las propias restricciones impuestas por el sistema cognitivo para su uso. Como hemos visto, si la cultura formatea la mente, también la mente formatea la cultura, ya que ambas, mente y cultura, están construidas no con unidades de información sin contenido, con trozos de *plastilina*, sino con piezas de algún modo ya formateadas, como esos bloques del *lego*, las representaciones implícitas de la mente encarnada. Retomando la epidemiología representacional de Sperber (1996), según la cual las representaciones culturales se difunden o transmiten socialmente mediante mecanismos de *contagio cognitivo*, podemos decir que nuestro equipamiento cognitivo de serie actúa de algún modo como un sistema cognitivo inmunológico que nos previene contra algunos de esos contagios, de modo que ciertas representaciones culturales difícilmente se generalizan como representaciones mentales explícitas, ya que su posible adquisición debe enfrentar las restricciones impuestas por la propia estructura cognitiva de la mente. Las dificultades para generalizar o extender entre nosotros ciertas representaciones culturales, como por ejemplo el conocimiento científico (p. ej., Pozo y Gómez Crespo, 1998), son un ejemplo de estas restricciones impuestas por la mente a las representaciones culturales. Por más que la ciencia lleve siglos manejando diversas representaciones de la nada y el vacío, como muestra Barrow (2000), la mayor parte de las personas tienen serias dificultades para hacer que esas representaciones culturales «corran» en su mente, y en su lugar se aferran a sus representaciones encarnadas de la materia (Pozo y Gómez Crespo, 2002).

Dado que la función última de los conocimientos transmitidos culturalmente es optimizar la eficiencia cognitiva de la mente humana en determinados contextos sociales, el estudio de cómo han evolucionado algunos de esos sistemas culturales de conocimiento, como la escritura (Olson, 1994), los sistemas de numeración (Barrow, 2000; Bishop, 1991; Nunes y Bryant, 1997), los sistemas de representación gráfica (Postigo y Pozo, 1999, 2000; Wainer y Velleman, 2001), o la representación del tiempo (Pozo, 2001), puede ayudarnos a entender algunas de esas restricciones cognitivas que han condicionado su desarrollo. En este sentido deberíamos diferenciar entre cómo se han generado culturalmente esos sistemas, de forma explícita y lenta, a lo largo de generaciones, y cómo la mente humana los adquiere posteriormente, cómo se apropia de ellos por procesos tanto explícitos como implícitos y en un periodo de tiempo condensado (de Vega, 1995). Así, desde el punto de vista cognitivo al menos, es posible que muchas formas culturales de conocimiento, que como tales han surgido de forma explícita y deliberada, a través de una evolución lenta y gradual, en el momento de su transmisión cultural se «condensen» en aprendizajes en gran medida implícitos, de forma que pierden buena parte de sus significados explícitos originales. Estudiar estas formas de «condensación cognitiva» de los sistemas culturales de conocimiento es en mi opinión uno de los grandes retos para la investigación en psicología del conocimiento si queremos, finalmente, encontrar en ella a Hal. Para ello debemos estudiar no sólo los procesos de aprendizaje e instrucción específicos que hacen posible la adquisición de esos conocimientos culturales sino también el modo en que esos sistemas culturales de conocimiento, en su propia evolución histórica, se adaptan a las restricciones cognitivas impuestas por la mente humana, haciendo más fácil su adquisición y transmisión social.

Es en esa construcción mutuamente restringida entre la mente humana y la cultura, en esa «co-construcción» (Valsiner, 1996) entre los sistemas culturales y cognitivos de conocimiento, donde tal vez debamos buscar a Hal, ese Hal que habita en cada uno de nosotros y que, a pesar de sus limitadas capacidades de cómputo, de su célebre «racionalidad limitada», o tal vez, gracias a ella, dispone de competencias y recursos cognitivos que ningún otro sistema cognitivo, incluido hoy por hoy el sistema de conocimientos generado por la propia psicología cognitiva, logra no ya emular sino ni siquiera comprender.

REFERENCIAS

- Anderson, J.R. (1996). Implicit memory and metacognition: why is the glass half full? En L.M. Reder (Ed.), *Implicit memory and metacognition*. Mahwah, NJ: Erlbaum.
- Arsuaga, J.L. (2001). *El enigma de la esfinge*. Barcelona: Plaza y Janés.
- Baddeley, A. (1997). *Human memory. Theory and practice. 2nd edition*. Londres: Taylor And Francis. Trad. cast. de G. Evangelista: *Memoria humana*. Madrid: McGraw Hill, 1999.
- Barrow, J. (2000). *The book of nothing*. Londres: Jonathan Cape. Trad. cast. de J. García Sanz: *El libro de la nada*. Barcelona, Crítica, 2001.
- Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-660.
- Bishop, A. (1991). *Mathematical enculturation*. Dordrecht, Holanda: Kluwer Academic Press. Trad. cast. de G. Sánchez: *Enculturación matemática*. Barcelona: Paidós, 1999.
- Bruner, J. (1983). *In search of mind*. New York: Harper & Row. Trad. cast. de J.J. Utrilla: *En busca de la mente*. México DF.: FCE., 1985.

- Bruner, J. (1990). *Acts of meaning*. Harvard: Presidente and Fellows of Harvard college. Trad. cast. de J.C. Gómez Crespo y J. Linaza: *Actos de significado. Más allá de la revolución cognitiva*. Madrid: Alianza, 1991.
- Bruner, J. (1993). Do we acquire culture or vice versa? *Behavioral and Brain Sciences*, 16 (3), 515-516.
- Cairns-Smith, A.G. (1996). *Evolving the mind: on the nature of matter and the origin of consciousness*. Cambridge, UK: Cambridge University Press. Trad. cast. de M. González: *La evolución de la mente*. Madrid: Cambridge University Press, 2000.
- Caparrós, A. y Gabucio, F. (1986). La aparición del paradigma cognitivo: una cuestión problemática. *Revista de Historia de la Psicología*, 7 (2), 53-58.
- Carretero, M. (1997). *Introducción a la psicología cognitiva*. Buenos Aires: Aique.
- Carretero, M., Pope, M., Simmons, P.J. & Pozo, J.I. (Eds.) (1991). *Learning and Instruction. European Research in an International Context*. London: Pergamon Press.
- Clark, A. (1997). *Being there. Putting brain, body and world together again*. Cambridge, Mass: The MIT Press. Trad. cast. de G. Sánchez: *Estar ahí. Cerebro, cuerpo y mundo en la nueva ciencia cognitiva*. Barcelona: Paidós, 1999.
- Cole, M. (1996). *Cultural psychology. A once and future discipline*. Cambridge, Mass.: Harvard University Press. Trad. cast. de T. del Amo: *Psicología cultural. Una disciplina del pasado y del futuro*. Madrid: Morata, 1999.
- Damasio, A. (1994). *Descartes's error. Emotion, reason and the human brain*. New York: Avon Books. Trad. cast. de J. Ros: *El error de Descartes*. Barcelona: Crítica, 1996.
- Damasio, A. (1999). *The feeling of what happens. Body and emotion in the making of consciousness*. New York: Harcourt Brace and Company. Trad. cast. de F. Páez de la Cadena: *La sensación de lo que ocurre*. Madrid: Debate, 2001.
- De Vega, M. (1981). Una exploración de los metapostulados de la psicología cognitiva. *Análisis y modificación de conducta*, 16, 345-376.
- De Vega, M. (1982). La metáfora del ordenador: implicaciones y límites. En I. Delclaux y J. Seoane (Eds.), *Psicología cognitiva y procesamiento de la información*. Madrid: Pirámide.
- De Vega, M. (1995). Representaciones mentales: paradojas, debates y soluciones. En J. Mayor y M. de Vega (Eds.), *Memoria y representación*. Madrid: Alhambra.
- De Vega, M. (1998). La psicología cognitiva: ensayo sobre un paradigma en transformación. *Anuario de Psicología*, 29 (2), 21-44.
- De Vega, M. (2002). Del significado simbólico al significado corpóreo. *Estudios de Psicología*, 23 (2), 153-174.
- Denis, M. (1991). *Image and cognition*. New York: Harvester-Wheatsheaf.
- Dienes, Z. & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, 22, 735-808.
- Donald, M. (1991). *Origins of the modern mind. Three stages in the evolution of culture and cognition*. Cambridge, MA: Harvard University Press.
- Donald, M. (2001). *A mind so rare. The evolution of human consciousness*. New York: Norton.
- Edelman, G. (1987). *Neural Darwinism*. New York: Basic Books.
- Flavell, J.H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology*, 50, 21-45.
- Fodor, J.A. (1983). *The modularity of mind*. Cambridge, Mass: The MIT Press. Trad. cast. de J.M. Igoa: *La modularidad de la mente*. Madrid: Morata, 1986.
- Froufe, M. (1996). *El inconsciente cognitivo*. Madrid: Ediciones de la UAM.
- Gardner, H. (1985). *The mind's new science*. New York: Basic Books. Trad. cast: *La nueva ciencia de la mente*. Barcelona: Paidós.
- Glenberg, A. (1997). What memory is for. *Behavioral and Brain Sciences*, 20, 1-55.
- Gómez Crespo, J.C. (1998). La psicología del significado. *Anuario de Psicología*, 29 (2), 185-186.
- Gómez Crespo, J.C. y Núñez, M. (1998). Introducción: La mente social y la mente física: desarrollo y dominios de conocimiento. *Infancia y Aprendizaje*, 84, 5-32.
- Gómez Crespo, M.A. y Pozo, J.I. (2001). La consistencia de las teorías sobre la naturaleza de la materia: una comparación entre las teorías científicas y las teorías implícitas. *Infancia y Aprendizaje*, 24 (4), 441-459.
- Gopnik, A. y Meltzoff, A.N. (1997). *Words, thoughts and theories*. Cambridge MA: MIT Press. Trad. cast. de M. Sotillo e I.S. Wildschütz: *Palabras, pensamientos y teorías*. Madrid: Visor, 1999.
- Gould, S. (1991). Exaptation: a crucial tool for an evolutionary psychology. *Journal of Social Issues*, 47 (3), 43-65.
- Hirschfeld, L. & Gelman, S. (Eds) (1994). *Mapping the mind*. Cambridge, MA: Cambridge University Press.
- Holland, J. M., Holyoak, K. J., Nisbett, R. E. & Thagard, P. R. (1986). *Induction. Processes of inference, learning and discovery*. Cambridge, Mass: The MIT Press.

- Humphrey, N. (1986). *The inner eye*. Oxford: Faber and Faber. Trad. cast. de M.V. Laa: *La mirada interior*. Madrid: Alianza, 1995.
- Humphrey, N. (1992). *A history of the mind: Evolution and the birth of consciousness*. New York: Simon and Schuster. Trad. cast: *Una historia de la mente*. Barcelona: Paidós.
- James, W. (1890). *The Principles of Psychology*. New York: Holt.
- Karmiloff-Smith, A. (1992). *Beyond modularity*. Cambridge, Mass.: Cambridge University Press. Trad. cast. de J.C. Gómez y M. Núñez: *Más allá de la modularidad*. Madrid: Alianza, 1994.
- Kintsch, W. (1998). *Comprehension. A paradigm for cognition*. Cambridge Mass.: Cambridge University Press.
- Kirsh, D. & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18, 513-549.
- Kirsner, K. et al. (Ed.) (1998). *Implicit and explicit mental processes*. Mahwah, NJ: Erlbaum.
- Leslie, A. (1987). Pretense and representation: The origins of «theory of mind». *Psychological Review*, 94, 412-426.
- Lewontin, M. (2000). *It isn't necessarily so: The dream of human genome and other illusions*. New York: The New York Review of Books. Trad. cast. de R. Ibero: *El sueño del genoma humano y otras ilusiones*. Barcelona: Paidós, 2001.
- Martí, E. (Comp.) (1997). *Construir una mente*. Barcelona: Paidós.
- Martí, E. y Pozo, J.I. (2000). Más allá de las representaciones mentales: la adquisición de los sistemas externos de representación. *Infancia y Aprendizaje*, 90, 11-30.
- Mateos, M. (1995). *Mente y computación*. Madrid: Ediciones de la UAM.
- Neal, A. & Hesketh, B. (1997). Episodic knowledge and implicit learning. *Psychonomic Bulletin and Review*, 4 (1), 24-37.
- Nelson, K. (1996). *Language in cognitive development*. The emergence of the mediated mind. Cambridge: Cambridge University Press.
- Nisbett, R. (Ed.) (1993). *Rules of reasoning*. Hillsdale, NJ: Erlbaum.
- Nisbett, R., Peng, K., Choi, I. & Norenzayan, A. (2001). Culture and systems of thought: holistic vs. analytic cognition. *Psychological Review*, 108 (2), 291-310.
- Nunes, T. & Bryant, P. (Eds.) (1997). *Learning and teaching Mathematics. An international perspective*. London: Guilford Press.
- Olson, D. (1994). *The world on paper*. Cambridge: Cambridge University Press. Trad. cast. de P. Wilson: *El mundo sobre el papel*. Barcelona: Gedisa, 1999.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge MA: MIT Press. Trad. cast. de M.A. Galmarini: *Comprender la mente representacional*. Barcelona: Paidós, 1994.
- Pinker, S. (1997). *How the mind works*. New York: Norton. Trad. Cast.: *Cómo funciona la mente*. Madrid: Debate, 2001.
- Postigo, Y. y Pozo, J.I. (1999). Hacia una nueva alfabetización: el aprendizaje de información gráfica. En J.I. Pozo y C. Monereo (Eds.), *El aprendizaje estratégico: enseñar a aprender desde el currículo*. Madrid: Santillana/Aula XXI.
- Postigo, Y. y Pozo, J.I. (2000). Cuando una gráfica vale más que 1000 datos: la interpretación de gráficas por alumnos adolescentes. *Infancia y Aprendizaje*, 90, 89-110.
- Pozo, J.I. (1988). De las tormentosas relaciones entre forma y contenido en el razonamiento: crónica de un romance anunciado. *Estudios de Psicología*, 35, 117-135.
- Pozo, J.I. (1989). *Teorías cognitivas del aprendizaje*. Madrid: Morata.
- Pozo, J.I. (1997). El cambio sobre el cambio: hacia una nueva concepción del cambio conceptual en la construcción del conocimiento científico. En M.J. Rodrigo y J. Aray (Eds.), *La construcción del conocimiento escolar*. Barcelona: Paidós.
- Pozo, J.I. (2001). *Humana mente: el mundo, la conciencia y la carne*. Madrid: Morata.
- Pozo, J.I. y Gómez Crespo, M.A. (1998). *Aprender y enseñar ciencia. Del conocimiento cotidiano al conocimiento científico*. Madrid: Morata.
- Pozo, J.I. y Gómez Crespo, M.A. (2002). Más allá del «equipamiento cognitivo de serie»: la comprensión de la naturaleza de la materia. En M. Benlloch (Ed.), *La educación en ciencias: ideas para mejorar su práctica*. Barcelona: Paidós.
- Pozo, J.I., Gómez Crespo, M.A. & Sanz, A. (1999). When conceptual change does not mean replacement: Different representations for different contexts. En W. Schnotz; S. Vosniadou & M. Carretero (Eds.), *New perspectives on conceptual change*. London: Elsevier.
- Pozo, J.I. y Rodrigo, M.J. (2001). Del cambio de contenido al cambio representacional en el conocimiento conceptual. *Infancia y Aprendizaje*, 24 (4), 407-423.
- Reber, A.S. (1993). *Implicit learning and tacit knowledge*. New York: Oxford University Press.
- Reber, A.S. (1995). *Dictionary of Psychology*. Oxford: Penguin Books.

- Reigeluth, Ch.M. (Ed.) (1999). *Instructional-design theories and models. A new paradigm of instructional theory*. Mahwah, NJ: Erlbaum. Trad. cast. de E. Llavori y E. Alberola. *Diseño de la instrucción. Teorías y modelos* (2 vols.). Madrid: Santillana, 2000.
- Rivière, Á. (1987). *El sujeto de la psicología cognitiva*. Madrid: Alianza.
- Rivière, Á. (1991). *Objetos con mente*. Madrid: Alianza.
- Rivière, Á. (1997). Teoría de la mente y meta-representación. Informe no publicado. Facultad de Psicología, Universidad Autónoma de Madrid.
- Rivière, Á. y Núñez, M. (1996). *La mirada mental*. Buenos Aires: Aique.
- Rodrigo, M.J. y Correa, N. (2001). Representación y procesos cognitivos: esquemas y modelos mentales. En C. Coll, J. Palacios y A. Marchesi (Eds.), *Desarrollo psicológico y educación*. Vol II. Segunda edición: Psicología de la Educación escolar. Madrid: Alianza.
- Rosetti, Y. y Revonsuo, A. (Eds.) (2000). *Beyond dissociation: Interaction between dissociated implicit and explicit processing*. Amsterdam: John Benjamins.
- Russell, J. (1984). *Explaining mental life. Some philosophical issues in psychology*. London: Macmillan.
- Sacks, O. (1985). *The man who mistook his wife for a hat*. New York: Summit Books. Trad. cast. de J.M. Álvarez: *El hombre que confundió a su mujer con un sombrero*. Barcelona: Muchnick Editores.
- Salthouse, T. (1991). Expertise as the circumvention of human processing limitations. En K.A. Ericsson y J. Smith (Eds.), *Toward a theory of expertise*. Cambridge, Mass.: Cambridge University Press.
- Shanon, B. (1993). *The representational and the presentational. An essay on cognition and the study of mind*. London: Harvester.
- Spelke, E. & Newport, E. (1997). Nativism, empiricism and the development of knowledge. En R. Lerner (Ed.), *Handbook of child psychology*. Vol 1. New York: Wiley.
- Spelke, E., Phillips, A. & Woodward, A.L. (1995). Infant's knowledge of object motion and human action. En D. Sperber, D. Premack & A.J. Premack (Eds.), *Causal cognition. A multidisciplinary debate*. Oxford: Clarendon Press.
- Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Oxford: Blackwell.
- Sperber, D. (Ed.) (2000). *Metarepresentations. A multidisciplinary debate*. New York: Oxford University Press.
- Sperber, D. Premack, D. & Premack, A.J. (Eds.) (1995). *Causal cognition. A multidisciplinary debate*. Oxford: Clarendon Press.
- Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, Mass.: Harvard University Press.
- Tomasello, M. (2000). Primate cognition: Introduction to the Issue. *Cognitive Science*, 24 (3), 351-361.
- Tomasello, M., Krugcr, A.C. & Ratner, H. (1993). Cultural learning. *Brain and Behavioral Sciences*, 16 (3), 495-511.
- Tulving, E. (1989). Memory: performance, knowledge, and experience. *European Journal of Cognitive Psychology*, 1 (1), 3-26.
- Underwood, G. (Ed.) (1996). *Implicit cognition*. Oxford, Oxford University Press.
- Valsiner, J. (1996). Co-constructionism and development: A sociohistoric tradition. *Anuario de Psicología*, 69, 63-82.
- Valsiner, J. (1997). *Culture and the development of children's action*. 2nd ed. New York: John Wiley.
- Voss, J.F., Wiley, J. & Carretero, M. (1995). Acquiring intellectual skills. *Annual Review of Psychology*, 46, 155-181.
- Wellman, H. & Gelman, S. (1997). Knowledge acquisition in foundational domains. En D. Kuhn y R.Siegler (Eds.), *Handbook of child psychology*. Vol 2. New York: Wiley.
- Woll, S. (2002). *Everyday thinking*. Mahwah, NJ: Erlbaum.