

## Orígenes históricos de la psicología cognitiva: paradigma simbólico y procesamiento de la información

Ángel Rivière  
Universidad Autónoma de Madrid

*La psicología cognitiva supone la recuperación de una vieja tradición epistemológica en psicología. Sin embargo, implica un enfoque nuevo del supuesto de que gran parte de la conducta puede explicarse en virtud de representaciones internas e intencionales. La aproximación cognitiva es nueva por dos razones: 1) Explica mediante un vocabulario intencional observaciones establecidas en términos extensionales, y 2) incluye los elementos de ese vocabulario en nociones de cómputo, que implican el compromiso con un mecanicismo abstracto y formal. Los modelos sintácticos, formales y racionales predominaron en la primera fase de la historia de la psicología cognitiva. En los años setenta y ochenta, muchos investigadores respondieron a la necesidad de dar cuenta de las influencias cognitivas de los contenidos y de ofrecer modelos capaces de explicar las limitaciones y sesgos del procesamiento de la información. En la década de los años ochenta, el desarrollo de los modelos conexionistas implicó la elaboración de un lenguaje extensional para explicar la mente como sistema de cómputo.*

Palabras clave: Psicología cognitiva, paradigma simbólico, procesamiento de la información, conexionismo, Máquina de Turing.

*Cognitive psychology means the recovery of an old epistemological tradition in psychology. However, it implies a new approach of the assumption that behaviour—or a great deal of it—can be explained by internal and intentional representations. The cognitive approach is new because of two reasons: 1) It explains observations in extensional terms with intentional lexicon and 2) it includes this lexicon elements in computational terms, which imply a bind with abstract and formal mechanism. Syntactic models, formal and rational, prevailed during the first period of the history of cognitive psychology. In the 70's many researchers provided for information on cognitive influences of mental contents and offered models which could explain the limitations and biases of information processing. In the 80's, the development of connectionist models implied the working out of an extensional language to explain the mind as a computational system.*

Key words: Cognitive Psychology, Symbolic Paradigm, Information Processing, Connectionism, Turing's Machine.

## La tradición epistemológica de la psicología

Hace algunos años, Jerry Fodor (1988), uno de los defensores más lúcidos de la ortodoxia de los cómputos y los símbolos en psicología cognitiva, comentaba un hecho evidente: las personas empleamos, en nuestras interacciones ordinarias, una psicología «de sentido común», cuyo núcleo consiste en la atribución recíproca de creencias y deseos. Para los propósitos cotidianos de relación, esa psicofilia mentalista, que se basa en el supuesto cartesiano de que «las personas obran por conocimiento», es tan perfecta como un Rolls Royce: «La psicología del sentido común funciona tan bien que termina por desaparecer. Es como esos míticos coches de Rolls Royce, cuyos engranajes quedan sellados al salir de fábrica; pero es incluso mejor porque no es mítica» (p. 3). En realidad, la psicología mentalista, y en cierto sentido «cognitiva», del sentido común es una herramienta de comprensión y predicción psicosocial tan poderosa y eficaz que «si se colapsara, ésta sería, sin comparación, la mayor catástrofe intelectual en la historia de nuestra especie; si estuviéramos equivocados acerca de la mente, nuestra equivocación sería la mayor que hemos tenido nunca acerca de algo. Las consecuencias del colapso de lo supranatural, por ejemplo, no podrían compararse con éstas; el teísmo nunca ha estado tan encarnado en nuestro pensamiento y nuestras prácticas —especialmente en nuestras prácticas— como lo están las explicaciones en términos de creencias y deseos» (1988, p. 12).

Sin embargo, durante buena parte de nuestro siglo, muchos psicólogos han especulado seriamente con la posibilidad de producir, en el ámbito reducido de la psicología científica, ese supuesto colapso intelectual; con la idea de romper, en un sentido muy radical, con la vieja psicología del sentido común y con su noción esencial de que las personas obran en virtud del conocimiento o, lo que es lo mismo, en gracia de ciertas representaciones internas, semánticamente relacionadas con el mundo, y de naturaleza mental. Es posible que ese ideal, tan acariciado por los conductistas, de ruptura completa de la ciencia psicológica con respecto a las nociones implícitas más fundamentales y entrañadas de la psicología natural (nociones tan básicas como la de que la gente «hace en función de lo que conoce», «conoce mediante representaciones internas» y «tales entidades son de carácter intencional») ...es posible, sí, que esa pretensión de ruptura demuestre finalmente ser uno más de los monstruos de la razón de nuestro siglo atormentado. Al fin y al cabo, estamos en un periodo histórico en que algunos de los ideales más aventurados y ambiciosos nacidos en el propio siglo —o finales del pasado— se están derrumbando de forma rápida. Es cierto que todavía no ha llegado el momento en que podamos considerar al conductismo como un cachivache polvoriento arrumbado en una esquina de la historia. Sigue vivo, sobre todo, en la psicología aplicada. Pero sí podemos empezar nuestra propia historia —la historia de la psicología cognitiva— considerando que el conductismo constituyó, cuando menos, una importante *anomalía histórica* en el largo pasado y la breve historia de la psicología.

¿Por qué fue el conductismo una *anomalía histórica*? Para explicarlo, necesitamos alejarnos algo de la sombra que nos hace la psicología más inmediata (lo que resulta difícil ante la tarea de historiar algo tan presente como el enfoque

cognitivo) y tratar de hacer historia desde un horizonte más dilatado, con esa larga vista de pájaro que es la mirada propia de la historia. Desde esa visión lejana, el incidente del conductismo se nos muestra como un minúsculo episodio de ruptura epistemológica con una larguísima tradición en que la propia *episteme*, el conocimiento, había sido un tema esencial de cualquier indagación psicológica. La entrañable vinculación de la materia psicológica a la actividad de conocer no sólo se muestra en la obvia propensión epistemofílica de la psicología natural del sentido común (esa psicología intencionalista de creencias-deseos, que probablemente se originó en los primeros bocetos de la filogénesis de nuestra especie, y cuyo valor reivindica tan brillantemente Dennett, 1987), sino también en las primeras especulaciones filosóficas que tematizaron el concepto de *psiqué*. En el *Fedón*, por ejemplo, Platón sigue el siguiente curso de pensamiento, de orientación netamente epistemológica: puesto que cosas tales como la justicia y la verdad (1) *son* algo, (2) *no se ven* con los ojos (3) ni se perciben con los sentidos, (4) son la esencia de las cosas, y (5) sólo se conocen con el pensamiento, que (6) debe desprenderse por entero del cuerpo para conocer la auténtica realidad, (7) ésta se identifica con la esencia y (8) es de carácter ideal. Por consiguiente, (9) en el hombre debe haber una realidad no reductible al cuerpo (10) y radicalmente separable de él, (11) a la que corresponde el pensamiento, y es capaz de re-conocer los objetos ideales: la *psiqué*.

Si en el origen filosófico del primer objeto psicológico (la *psiqué* misma) jugaron un papel esencial las reflexiones epistemológicas —tales como las relacionadas con la inmaterialidad del conocimiento y la posibilidad de construir, descubrir o recordar objetos ideales con el pensamiento—, los temas epistemológicos pasaron a ocupar un puesto aún más central en el momento de diferenciación de la mente como objeto de reflexión, en el pensamiento moderno. Así, en la filosofía racionalista de Descartes, la mente se define como la sustancia que piensa y se subraya el papel epistemológico del pensamiento en su capacidad de concebir y deducir objetos lógicos y matemáticos que habrían de configurar, a largo plazo, las bases de cimentación de la ciencia cognitiva como un todo (no debemos olvidar que los conceptos de «algoritmo» y «cómputo», tienen su entronque en el ideal racionalista, de Descartes y Leibniz, de un lenguaje automático y completo para el razonamiento). Por otra parte, la versión empirista del conocimiento acentuó aún más la estrecha relación entre los objetos psicológicos y los cognoscitivos, puesto que implicó una reducción esencial de la epistemología a psicología.

La larga tradición de reflexión epistemológica fue incorporada a la primera psicología «científica» de los grandes fundadores, como Fechner, Wundt, James y Brentano. Boring (1978), por ejemplo, destaca cómo la psicología experimental de Wundt fue, en cierta medida, un producto del mestizaje de la epistemología empirista de la tradición asociacionista con el enfoque metodológico de la investigación fisiológica alemana del siglo XIX; de manera que no es del todo inexacto decir que Wundt era un epistemólogo de bata blanca (los factores sociales que determinaron esa hibridación han sido analizados en un interesante artículo ya clásico de Ben-David y Collins, 1966). Esa misma pasión epistemológica guía las investigaciones de Külpe, Buhler, Wertheimer, Köhler, Vygotski,

Piaget y Bartlett, y su rescoldo nunca se apagó del todo en la psicología europea, ni siquiera en los años de predominio indiscutible del enfoque «epistemofóbico» en la psicología objetivista del conductismo norteamericano. Por todo ello, podemos decir que la psicología *volvió a ser* cognitiva en la segunda mitad de nuestro siglo, y de ninguna manera que *empezó a ser cognitiva* en los años cincuenta y sesenta. El enfoque cognitivo en psicología ha supuesto la recuperación explícita de la viejísima tradición epistemológica de la psicología natural de sentido común, de la reflexión filosófica sobre el alma, la mente y la conciencia, y de la primera psicología científica. Se enraiza en lo más hondo y viejo de la historia de nuestra disciplina, pero lo hace de un modo peculiar. La psicología cognitiva lo es de un modo diferente a la forma en que fueron «cognitivas» esas otras psicologías.

Para explicar este galimatías histórico —el rompecabezas de «una psicología que volvió a ser cognitiva después de dejar de serlo de forma provisional y local, pero siéndolo de distinto modo a como lo había sido antes»— conviene que analicemos brevemente el incidente del conductismo, y que nos detengamos en sus justificaciones y en las que existen para decir, como hemos dicho sin ánimo ofensivo, que el enfoque conductista era «epistemofóbico». El problema fundacional del conductismo fue, como bien sabe cualquier psicólogo, el de la *objetividad*. ¿Cómo hacer de la psicología una ciencia objetiva? En 1913, Watson ligaba explícitamente ese problema a lo que podríamos llamar el escándalo de la introspección, la conciencia y los objetos mentales en general.

El escándalo consistía en que, a diferencia de lo que sucede con los objetos sobre los que versan las ciencias objetivas de la naturaleza, esas otras «entidades aéreas» (de las que habla ingenuamente la psicología epistemofílica del sentido común cuando dice que la gente tiene cosas tales como imágenes, representaciones mentales, pensamientos, etc.) parecen inevitablemente encadenadas al oscuro reducto de la conciencia y ésta, a su vez, al más que dudoso recurso observacional de la introspección. La propuesta de Watson era clara, ¿por qué no hacer de lo que podemos *observar* y *medir* el verdadero objeto de la psicología?, ¿por qué no eliminar un método que es *observacionalmente asimétrico*, vinculado a la primera persona de singular (la que *se* observa a sí misma), y por todo ello inútil para hacer de la psicología una ciencia?... y, lo que parecía seguirse de todo ello, ¿por qué no abandonar del todo, o traducir a otro lenguaje las etéreas entidades, vinculadas a ese método, de la psicología natural y de la primera psicología introspectiva?

A pesar de que existía una tradición histórica que había roto la ecuación mente-conciencia, y que se remontaba a figuras venerables como las de Leibniz, Herbart y Freud, por el año 1913 resultaba difícil diferenciar con nitidez los conceptos de mente y conciencia y comprender cómo podría desarrollarse una ciencia *objetiva* acerca de la *mente* que no pasara por ese lugar oscuro, impreciso, de la conciencia. Mente y objetividad parecían sustancias tan incompatibles como el fuego y el agua.

Los objetos incluidos en el vocabulario teórico y en la base observacional de la psicología introspectiva, como los de la psicología del sentido común, poseen dos propiedades que parecen invalidarlos como objetos de indagación cien-

tífica: son *internos* y son *intencionales* (Brentano, 1874), es decir, implican *representaciones* acerca de algo que no se identifica con las propias representaciones y a primera vista sólo parecen realmente accesibles por experiencia interna. En su naturaleza interna y su carácter intencional residen sus posibilidades epistémicas. Para los ingenuos psicólogos naturales y los sesudos introspeccionistas que pretendieron ser, a la vez, científicos y psicólogos, las personas son, efectivamente, seres que, como decía Descartes, «obran por conocimiento y no por la disposición de sus órganos». Esa capacidad de «obrar por conocimiento» era la que separaba, para Descartes, a las personas de los animales, y su negación en éstos abrió la vía al programa mecanicista de interpretación de la conducta animal que heredaron los objetivismos reflexológico y conductista.

Imaginar *algo*, pensar sobre *algo*, recordar *algo*, son actividades epistémicas e intencionales —como decía Brentano—, mientras que responder a energías físicas no lo es. En esta perspectiva, el programa conductista consistía, en esencia, en *hacer una psicología cuyos enunciados teóricos y empíricos fueran escuetamente extensionales*, es decir, potencialmente exteriorizables a una observación en tercera persona (y no en primera) y no consistentes en representaciones *acerca de*, sino en *respuestas a*. En la obra de Watson se muestra esta pretensión con la mayor claridad: resulta muy reveladora, por ejemplo, su intención de reducir el pensamiento a *habla subvocal*. ¡Claro!... la acción de *hablar* no es, en sí misma, una actividad *intencional* —y, sin embargo, sí lo es la de decir *algo*— y el verbo *hablar* expresa el componente más periférico, externo y «desintencionado» de la acción de decir. La formulación watsoniana final de que «pensamos con todo el cuerpo» es la explicitación más clara de una concepción de la psicología ingenuamente «realista» y que establecía unos *criterios de justificación* de las entidades psicológicas muy ajenos a los de la psicología introspectiva primera.

¿Cuáles eran los nuevos criterios para las entidades psicológicas? Esencialmente dos: (1) que tales entidades fuesen real o potencialmente *observables* (como por ejemplo lo es la acción de «hablar» y no la de «pensar»), (2) y que sus mecanismos de formación y transformación fuesen reductibles, en último término, a procesos asociativos simples de aprendizaje y condicionamiento. Bever, Fodor y Garrett (1968) han denominado «metapostulado terminal» del asociacionismo al siguiente conjunto de proposiciones, que definen la posición conductista: (1) *Proposición sensista*: los elementos requeridos por una explicación psicológica pueden situarse en correspondencia biunívoca con elementos potencialmente observables (estímulos, respuestas o derivados de ellos), (2) *Proposición conexionista*: dichos elementos se conectan o asocian entre sí sólo cuando se dan en contigüidad, (3) *Proposición mecanicista*: cualquier conducta observable puede explicarse por concatenación de tales vínculos asociativos.

La posibilidad de éxito de la ambiciosa operación de reducción objetivista del viejo vocabulario intencionalista de la psicología natural dependía de una reducción previa, a términos puramente *extensionales*, de las nociones de estímulo y respuesta. Como ha demostrado lúcidamente Yela (1974), esa pretensión, en la que se basaba toda la operación conceptual del conductismo, nunca pudo lograrse del todo, y las nociones de «estímulo» y «respuesta» eran muy ambiguas en Watson, que definía, por ejemplo, la primera en términos de «las energías fi-

sicas» que inciden sobre los sentidos (es decir, objetos claramente *extensionales*), unas veces, y como «aquello a lo que responde el organismo» (es decir, un objeto *intencional*), en otras ocasiones.

Las observaciones anteriores nos sirven para caracterizar muy brevemente la *novedad* de la nueva psicología cognitiva de la segunda mitad de nuestro siglo, con relación tanto a la vieja psicología introspectiva como al conductismo. Mientras que aquella pretendió, sin éxito, basarse en un vocabulario observacional *intencional* (el proporcionado por la introspección, que define objetos tales como «pensamientos», «imágenes», *bewusstseinslagen*, etc.) para desarrollar una psicología con un vocabulario teórico también interno e intencional, y éste —el conductismo— pretendió una psicología escuetamente *extensional*, tanto en su lenguaje observacional como en el teórico (compuesta de entidades *físicas* y reduciendo el problema del *significado* al de la reinstauración de respuestas o a la relación funcional entre éstas y las condiciones estímulares), la novedad de la psicología cognitiva ha consistido esencialmente en tratar de hacer una psicología que fundamentase un lenguaje *intencional* acerca de la mente en observaciones establecidas en términos estrictamente *extensionales*. Es eso, por ejemplo, lo que hace el psicólogo cognitivo que fundamenta sus conceptos de «proposición», «esquema», «rasgo semántico», en los tiempos de respuesta de sus sujetos en tal o cual tarea, o en los errores que cometen, etc. Los esquemas, las proposiciones y los rasgos semánticos son objetos internos e intencionales, es decir, entidades mentales de naturaleza epistémica. Los tiempos de respuesta y los errores admiten descripciones *extensionales*.

La explicación anterior es completamente equivalente a la que se ha propuesto tradicionalmente, cuando se ha dicho que la psicología cognitiva *trata de ser una ciencia objetiva de la mente*. Intenta resolver el viejo nudo gordiano que no sabían cómo desatar los psicólogos de comienzos de nuestro siglo: el de hacer compatibles fuego y agua, mente y objetividad. Pero el proceso histórico que condujo a la recuperación de la mente no fue sólo un episodio interno de la historia de la psicología. La interpretación histórica tradicional, según la cual la vuelta de lo mental fue principalmente un resultado de las anomalías e insuficiencias del conductismo y de una rebelión estrictamente interna contra la exigente dieta conceptual a la que obligó el intento objetivista de prescindir de los objetos mentales, presenta sólo una parte, y muy parcial, de la verdad. El origen histórico de la psicología cognitiva no puede entenderse si se analizan los hechos con la visión limitada de la *historia interna* de la psicología. Muy al contrario: representa el momento en que irrumpe con mayor fuerza, en psicología, la *historia externa*, exigiendo una nueva concepción de la mente que no era compatible ni con el reduccionismo *extensional* del conductismo ni con las vagas intuiciones mentalistas de la psicología introspectiva natural o científica.

En suma, para comprender los orígenes históricos de la psicología cognitiva no sólo es necesario remontarse, en el tiempo, a la vieja tradición epistemológica de la psicología, sino también salirse de la lógica interna de la historia de nuestra disciplina, analizando los factores externos que han dado lugar al desarrollo de un modo de hacer psicología que forma parte de un proyecto científico más general, el de la *ciencia cognitiva* como un todo, y que se relaciona con las

*tecnologías del conocimiento* que configuran quizá los utensilios que definen, de manera más peculiar, la tecnología propia de la segunda mitad de nuestro siglo. Con el desarrollo del enfoque cognitivo, la psicología ha pasado a formar parte de uno de los proyectos interdisciplinarios más ambiciosos de nuestro siglo y de ese conjunto privilegiado de ciencias que producen tecnología. Las consecuencias históricas de este hecho son enormes. ¿Cómo se ha producido ese peculiar proceso de maridaje entre mente —o intencionalidad— y tecnología? Frente a las visiones tradicionales —como la del propio Descartes— que contraponían cualquier visión mecanicista de la naturaleza al concepto de mente, la psicología cognitiva se ha basado en una forma especial de mecanicismo que parecía justificar, con nuevos bríos, la vieja pretensión de hacer una ciencia objetiva y rigurosa de la mente.

### La mente como máquina abstracta y el proyecto de la psicología cognitiva

Se dice normalmente —y nosotros acabamos de hacerlo— que el enfoque cognitivo supuso la restitución de los derechos de lo mental en el ámbito disciplinar de la psicología. Esta afirmación es muy cierta, pero puede resultar equívoca si no se cualifica. En realidad no fue la vieja idea de la mente la que resucitó con el enfoque cognitivo. La psicología cognitiva tiene una actitud tan escéptica como la conductista —aunque por otras razones— hacia esa huidiza *mente fenoménica*, que se ofrece a la introspección ingenua del psicólogo natural o a la más sistemática del introspeccionista avezado.

Una razón de la importancia histórica del enfoque cognitivo es que *implicó la delimitación y definición de un nuevo plano de lo mental, al considerar a la mente como un sistema de cómputo*. Como dice Johnson-Laird, «la invención del ordenador digital, y de manera más importante, de su precursora, la teoría matemática de la computabilidad, ha obligado a la gente a pensar de una forma nueva sobre la mente. Antes de la computación había una distinción clara entre cerebro y mente; uno era un órgano físico y la otra una «no entidad» fantasmática que difícilmente resultaba un tema de investigación respetable. (Se consentía que los adultos pudiesen hablar de ella en privado, siempre y cuando comprendieran que, en realidad, no existía.) Después de la llegada de los ordenadores no cabe semejante escepticismo: una máquina puede controlarse mediante un «programa» de instrucciones simbólicas, y no hay nada de fantasmal en un programa de ordenador. Quizá, y en gran medida, la mente es para el cerebro lo que el programa es para el ordenador. De esta manera, puede haber una ciencia de la mente» (1990, pp. 13-14).

Uno de los aspectos del origen de la psicología cognitiva que merece una interpretación histórica es éste: por qué la mente había dejado de ser respetable en 1913 y volvió a serlo en los años sesenta de nuestro siglo. Cómo tuvieron que vestirse esos peculiares y fantasmáticos miembros de su cohorte (los pensamientos, las imágenes, los recuerdos y creencias) para ser readmitidos en el recinto

respetable de la ciencia. La respuesta, como sugiere Johnson-Laird, es muy clara: tuvieron que presentarse como *objetos computables* o como *productos de algoritmos de cómputo*. Este ropaje sí era respetable: por una parte, desligaba los conceptos mentales de la dudosa fenomenología de la conciencia, además permitía simular su funcionamiento mediante algoritmos rigurosos que dieran cuenta de sus «productos extensionales» (en especial, de los tiempos de respuesta) y los revestía del prestigio de la tecnología emergente del conocimiento. Finalmente, los incluía en un contexto preciso de interpretación mecanicista, aunque se tratara de un mecanicismo muy diferente, mucho más complejo, elaborado y abstracto, que el empleado por los psicólogos objetivistas de la primera mitad del siglo.

La noción de cómputo se ligó intuitivamente a lo mental desde su nacimiento. Fue formulada rigurosamente, en 1936, por un matemático inglés excéntrico y genial, Alan Turing, que imaginó una máquina abstracta muy simple, capaz de simular cualquier actividad que pudiera descomponerse en un *procedimiento efectivo*, es decir, en un algoritmo consistente en una serie de transformaciones discretas de cadenas, también discretas, de símbolos. Es computable todo aquello que es realizable por tal máquina abstracta y universal. Turing intuyó enseñada las posibilidades que abría su máquina abstracta para una nueva comprensión de la mente, y planteó un desafío de consecuencias históricas aún incalculables: ¿por qué no romper con la tradicional división entre «mentes» y «máquinas», tan trabajosamente lograda por una ciencia de la naturaleza que había tardado siglos en deshacerse del todo de sus últimos residuos animistas?, ¿por qué no imaginar una máquina que produce prototipos de lo mental, es decir, «pensamientos»? Para la historia de la psicología, el desafío de Turing en «*Computing Machinery and Intelligence*» (1950) tenía una importancia enorme, porque parecía abrir perspectivas de solución a un viejo problema nunca resuelto: el de acomodar la *explicación mecanicista*, requerida por la necesidad de desarrollar una ciencia objetiva, con el concepto de *mente*. Hasta Turing, los intentos mecanicistas y objetivos (tales como los de la reflexología y el conductismo) se habían saldado inevitablemente con soluciones eliminacionistas: con la eliminación de lo mental o su reducción a términos puramente extensionales y, por consiguiente, no-mentales.

Turing (1950) proponía un juego estrictamente *objetivo* (muy conductista, en realidad) para provocar la reflexión y cuestionar las ideas tradicionales sobre la irreductible separación entre mentes y máquinas. Supongamos que estuviera encerrada, en una habitación, una «mente» (es decir, una persona) y una máquina en otra, y que alguien, quizá un psicólogo objetivista, hiciera preguntas a las dos sin saber quién respondía. Si ese «psicólogo objetivo» fuera incapaz de distinguir a la mente de la máquina, ¿con qué derecho podríamos decir que la primera piensa y la segunda no? Para Turing, si los «pensamientos» de las personas pueden descomponerse en *procedimientos efectivos* o algoritmos precisos, podrán ser simulados por una máquina universal como la ideada previamente por él mismo (Turing, 1936), capaz de computar cualquier procedimiento efectivo. Por otra parte, si el pensamiento *no* pudiera descomponerse en procedimientos efectivos, sería literalmente *incomprensible*. La metáfora de Turing de la «máquina que piensa» planteaba un desafío importante a la psicología, y abría nue-

vas posibilidades para ella. En realidad, la pregunta «¿pueden pensar las máquinas?» admitía una lectura propiamente psicológica: «Si las personas también piensan, ¿qué hay de misterioso en esta actividad?». En una palabra: la formulación de Turing era la primera expresión de lo que se ha llamado, en psicología cognitiva, «la metáfora del ordenador».

La máquina imitadora del pensamiento, ideada por Turing, no era una máquina energética, sino simbólica. En realidad, consistía en cadenas de unos y ceros que, con arreglo a los automatismos prefijados, serían capaces de computar cualquier algoritmo. La naturaleza *simbólica* de esa máquina que trataba información, su carácter abstracto e independiente de la sustancia material en que pudiera instanciarse, fue uno de los aspectos que, a la larga, habría de tener consecuencias más importantes para el proyecto primero de la psicología cognitiva. Permitía imaginar una alternativa mecanicista para lo mental, pero independiente de los aspectos extensionales y del substrato material en que lo mental se encarna. Por otra parte, la ilimitada universalidad de aquel autómatas, capaz de tratar cualquier procedimiento efectivo, constituía una propiedad importante: recordemos que Descartes, por ejemplo, afirmaba el carácter mecánico de los animales, y les negaba el atributo de pensar, basándose en que «aunque (los animales) hagan muchas cosas igual de bien, o quizá mejor que alguno de nosotros, carecerían infaliblemente de otras, *por lo que se descubriría que no obran por conocimiento, sino por la disposición de sus órganos*. Porque mientras que *la razón es un instrumento universal* que puede servir en toda clase de situaciones, esos órganos necesitan una disposición particular para cada acción particular, de donde se deriva que es moralmente imposible que haya suficientes en una máquina para hacerla obrar en todas las situaciones de nuestra vida, de la misma manera que nuestra razón nos hace obrar» (*Discurso del Método*, pp. 73-74 de ed. española de 1989);

¿Era una máquina, la de Turing, con «razón universal»? Sí, lo era en cierto modo, si la razón es algorítmica. Pero la condición de posibilidad para la nueva concepción mecanicista de la mente, que a la vez era aparentemente capaz de dar cuenta de ese carácter universal de la razón de que hablaba Descartes, fue despojar al concepto de «máquina» de toda connotación material. Con arreglo a la vieja distinción aristotélica, la máquina de Turing no es materia sino *forma*. Es una máquina formal y que, precisamente por su forma, es capaz de imitar a la mente.

Así, la psicología cognitiva se caracterizó, desde sus primeros esbozos, por un mecanicismo abstracto y por la tendencia a *ofrecer explicaciones formales de los fenómenos*, definiendo su génesis en virtud de la forma de los algoritmos que los producen. Esta posición, que ya se atisbaba en el artículo de Turing de 1950, estaba tan lejana del mecanicismo fisicista de Watson o Setchenov, como de la manera de explicar los fenómenos mentales en la vieja psicología introspectiva. Zenon Pylyshyn (1984) ha comentado, con mucha claridad, la importancia histórica que tenía la obra de Turing para la ciencia cognitiva como un todo, y para la psicología cognitiva en particular: «La obra de Turing —dice Pylyshyn— puede considerarse como el primer estudio de la actividad cognitiva, en que ésta se abstraía en principio por completo tanto de sus fundamentos biológicos como

de sus implicaciones fenomenológicas... representa la *emergencia de un nuevo nivel de análisis*, independiente del físico pero de espíritu mecanicista. Hace posible una ciencia de la estructura y la función divorciada de la sustancia material /.../ dado que habla el lenguaje de las estructuras mentales y los procesos internos, puede dar respuesta a cuestiones planteadas tradicionalmente por los psicólogos» (p. 68).

Esta misma caracterización es la que ofrece Howard Gardner, en una de las pocas aportaciones al análisis histórico de la ciencia cognitiva, *La nueva ciencia de la mente* (1987): esta nueva ciencia se caracteriza esencialmente por «concebir representaciones mentales y por postular un nivel de análisis totalmente separado del nivel biológico o neurológico» (p. 22). También se caracteriza por su adhesión a la metáfora del ordenador, es decir, a la noción de cómputo. Sin embargo, desde los comienzos de la ciencia cognitiva, hubo intentos de encarnar esa nueva versión formalista de la mente (la versión computacional) en modelos abstractos del sistema nervioso, saltando la barrera que separaba a la mente como mecanicismo simbólico, de la mente como función del cerebro. El más influyente de aquellos intentos fue el realizado por el neurofisiólogo Warren McCulloch y el matemático Walter Pitts, en un artículo de 1943, titulado «Un cálculo lógico inmanente en la actividad nerviosa» (reeditado en Boden, 1990). McCulloch y Pitts consideraban el cerebro como un sistema que incorpora principios lógicos en sus neuronas y redes neurales, defendiendo la idea de que todas las funciones computables por la máquina de Turing podrían ser realizadas por redes neurales, que serían capaces de tratar también todas las funciones del cálculo proposicional (de poder equivalente al de la máquina universal de Turing).

La relevancia histórica de la aportación de McCulloch y Pitts residía en que abría dos vías posibles de investigación para el nuevo enfoque objetivista de la mente que comenzaba a perfilarse por aquellos años: (1) una, centrada en la idea de *proposición*, entendida como herramienta simbólica de una mente descontextuada de su sustrato biológico, o *funcionalmente independiente* de él, (2) y otra que entendiera que el propio sistema nervioso puede concebirse, a nivel abstracto, como un sistema de cómputo, estableciendo para la psicología cognitiva un doble criterio de justificación: la adecuación a los datos comportamentales y la compatibilidad con los conocimientos neurobiológicos. Estos dos enfoques pueden identificarse, respectivamente, con el paradigma clásico C-R (es decir, de «Cómputos sobre Representaciones» simbólicas, en una máquina de carácter digital y esencialmente secuencial) y con la alternativa conexionista.

A pesar de las aparentes ventajas que hubiera podido tener la segunda versión —más «realista» en lo neurobiológico y que daba una cierta solución no dualista al problema de relación entre cuerpo y mente computacional—, fue la primera opción la que predominó en el enfoque cognitivo desde finales de los años sesenta a mediados de los ochenta, en que la opción «neural» fue retomada por el conexionismo. Ello se debió a dos factores: por una parte, el lenguaje intencionalista y simbólico de la alternativa clásica era mucho más directo y atractivo, para el psicólogo, que el más molecular y extensional de las redes neurales (de modo parecido a como el lenguaje de la física clásica es más directo e intuitivo, para el físico, que el de la mecánica cuántica). Por otra, los modelos de redes

desarrollados en los primeros años sesenta, y en concreto el «perceptrón» de Franz Rosenblatt (1962), compuestos por sólo dos capas de unidades, se demostraron pronto incapaces de realizar tareas simples tales como «comprender» la disyunción no-excluyente de la lógica de proposiciones (Minsky y Papert, 1969). El entusiasmo exagerado, y algo provocativo, de Franz Rosenblatt, el inventor del perceptrón, contribuyó bastante a que el proyecto de desarrollar modelos abstractos del sistema nervioso capaces de cómputo se desacreditara en cuanto aparecieron las primeras dificultades (McCorduck, 1991). Así, el camino quedó expedito para una versión de la psicología del conocimiento como ciencia objetiva de la mente, entendida esencialmente como sistema de cómputo de representaciones simbólicas, y situada en un plano funcionalmente independiente del nivel neurobiológico.

Esta alternativa, a pesar de revestirse desde muy pronto de una nueva forma de dualismo (el *dualismo funcionalista*), que hubiera resultado inaceptable para el paladar de cualquier psicólogo objetivista de los años cincuenta, contaba con el prestigio algorítmico de los cómputos y, en seguida, con la presencia real de los ordenadores, para justificar la nueva empresa de hacer una ciencia objetiva y rigurosa de la mente que superase las estrechas limitaciones conceptuales y metodológicas impuestas por el conductismo y el funcionalismo objetivista de los años cuarenta y cincuenta.

### El origen del paradigma simbólico-computacional de la psicología cognitiva

Pero, ¿significaba realmente la nueva psicología cognitiva, que apuntaba en los años cincuenta, una ruptura con el conductismo y sus criterios de justificación observacionales y teóricos?, ¿fue, de verdad, una «liberación»? Las interpretaciones tradicionales sobre el origen histórico de la psicología cognitiva han tendido a acentuar su vertiente revolucionaria con respecto al conductismo, y es cierto que un factor que influyó en su origen fue la percepción, por parte de muchos psicólogos, de que era demasiado lo que *no podía decirse* en el vocabulario conductista. Pero hay también otra versión de la historia: la que acentúa los rasgos de *continuidad* entre el paradigma cognitivo emergente y el enfoque conductista. Por ejemplo, Leahey presentó, en una reunión de la APA del año 1981, una ponencia con un título más bien provocativo: «La revolución que nunca existió: el procesamiento de la información *es* conductismo».

¿Es conductismo la nueva psicología cognitiva? Y, si lo es, ¿en qué sentido? De hecho hay rasgos de continuidad importantes: la mayoría de los psicólogos cognitivos han conservado la premisa conductista, según la cual *los enunciados observacionales de la psicología como ciencia deben ser extensionales y formulados en tercera persona de singular*, es decir, la exigencia del objetivismo de método. En áreas limitadas de estudio (como la solución de problemas) se han utilizado, desde algunos enfoques, informes depurados de naturaleza introspectiva (Newell y Simon, 1972; Ericsson y Simon, 1984), pero esta alternativa —que es la excepción más que la regla— suele compensarse con una acentuación de

las exigencias *sintácticas* y formalistas de los modelos teóricos, y corresponde a investigadores que no recogen tanto la tradición psicológica, propiamente dicha, como la de la Inteligencia Artificial (IA). También se produce una continuidad, mucho más matizada, en la pretensión mecanicista de los dos paradigmas, aunque se trate en un caso de un mecanicismo fisicista y en otro de un mecanicismo simbólico y abstracto.

Estas observaciones permiten definir los nuevos criterios de justificación que estableció el enfoque cognitivo para la psicología científica: podemos denominarlos (1) criterio de extensionalidad de la base empírica, y (2) criterio de justificación algorítmica de los conceptos teóricos. El primero se ha acentuado en la «versión débil» de la metáfora del ordenador, de propensión más empirista, y que subraya la importancia de la *base semántica* de las teorías cognitivas (De Vega, 1982). Es ésta la posición de la mayoría de los psicólogos empíricos que construyen teorías de la mente con el llamado «Enfoque de procesamiento de la información». Valorando más la exigencia de rigor observacional que la de rigor conceptual, y siguiendo la vieja tradición empirista de la psicología científica, justifican los resultados de sus experimentos en una analogía más bien vaga entre el ordenador y la mente, pero no están dispuestos a aceptar una versión literal del desafío de Turing. El criterio de justificación algorítmica es, en cambio, el que predomina en lo que podemos llamar «Paradigma computacional-representacional», en que se incluyen investigadores como Newell y Simon, Fodor o Pylyshyn. Son teóricos de propensión más racionalista que, asumiendo una versión fuerte y *literal* de la metáfora del ordenador, acentúan la importancia de la *forma sintáctica* de las teorías de la mente, aunque se sientan al tiempo más liberados de dar una justificación extensional de la base empírica (no es extraño así que Newell y Simon recurran a la introspección, o Chomsky a la «intuición del hablante» para justificar sus modelos de cómputos sobre representaciones).

Estos dos modelos han convivido, no siempre de buena gana, a lo largo del desarrollo de la psicología cognitiva. Un representante prototípico del paradigma C-R, Zenon Pylyshyn, comenta con disgusto que «a pesar del empleo generalizado de una terminología computacional (por ejemplo, términos como «almacenamiento», «proceso», «operación»), en gran parte de estos usos, los términos se emplean dándoles, como mucho, un contenido metafórico. Ha habido una resistencia a tomar la computación como una descripción *literal* de la actividad mental, y no sólo como una metáfora con un valor puramente heurístico... La falta de consideración de la tesis de la computación en un sentido literal ha posibilitado una amplia gama de actividades realizadas bajo la rúbrica de la «teoría del procesamiento de la información» algunas de las cuales representan una desviación significativa de las ideas que considero básicas en toda teoría computacional de la mente» (1980, p. 114).

Con independencia de esas importantes diferencias de matiz, que permiten defender la idea de que en la psicología cognitiva han convivido dos subparadigmas diferentes, lo cierto es que su caracterización histórica sólo puede hacerse tomando los dos criterios anteriores: modelos teóricos de propensión algorítmica y datos generalmente extensionales. La ruptura mentalista de la psicología cognitiva sólo puede entenderse cuando se sitúa en relación con esos dos criterios.

La recuperación de los constructos explicativos que implican o predicen conocimiento (esas entidades perdidas en el principio de la travesía conductista) sólo fue posible en el contexto de dos pretensiones científicamente respetables: (1) explicar la conducta, y (2) explicarla en términos mecánicos. Sólo así fue posible una tercera pretensión: (3) explicarla como producto del conocimiento. La recuperación de los viejos conceptos de la psicología natural mentalista, de los «planes» y propósitos (Miller, Galanter y Pribram, 1960), las estrategias activas de elaboración del conocimiento (Bruner, Goodnow y Austin, 1956), los símbolos mentales, como objetos portadores de conocimiento (Newell, Shaw y Simon, 1958), estuvo avalada por esos criterios de justificación.

Estas consideraciones, al tiempo que sirven para destacar los rasgos de continuidad entre el enfoque cognitivo y el conductista, ayudan a explicar el significado de las primeras críticas propiamente cognitivas de las explicaciones conductistas y el sentido de las primeras propuestas de una psicología del procesamiento de la información. Por ejemplo, en el importante simposio precursor celebrado en Hixon, en 1948, Karl Lashley explicó que los modelos explicativos conductistas «de izquierda a derecha» eran incapaces de dar cuenta de conductas organizadas, reclamando la necesidad de recurrir a organizaciones jerárquicas internas para explicar esas conductas. En contra de la tendencia anterior a reducir el lenguaje a conductas más simples, de naturaleza asociativa, Lashley reclamaba la necesidad de emplear un modelo gramatical o «lingüístico» —es decir, compo-nencial, estructurado y jerárquico— para explicar esas otras conductas motoras supuestamente más simples.

La posición de Lashley en la reunión de Hixon prefiguraba aspectos importantes de la que habría de ser la actitud cognitiva que se desarrollaría en años posteriores: la consideración del lenguaje como foco de interés, y de los modelos lingüísticos como modelos explicativos ejemplares, la insistencia en mecanismos de determinación «de arriba a abajo» y en sistemas jerárquicos de organización de la conducta, el gusto por la complejidad, iban a definir aquella nueva actitud, que se explicitó de forma mucho más completa ocho años después, en las Jornadas sobre Teoría de la Información, realizadas en 1956 en el *Massachusetts Institute of Technology* (MIT), institución que se configuró desde el primer momento como el «sancta sanctorum» de la nueva ciencia cognitiva.

La revisión de las comunicaciones presentadas en las Jornadas del MIT ofrece un balance muy preciso de las preocupaciones y posiciones que configuraban, por el año 1956, la nueva posición cognitiva. Un joven lingüista, Noam Chomsky, desarrollaba rigurosamente algunas intuiciones adelantadas por Lashley, al demostrar que las gramáticas de estados finitos y de izquierda a derecha, empleadas por los teóricos de la información, eran incapaces de generar el lenguaje. El modelo de gramática generativo-transformacional, que él proponía, se convirtió poco después en un prototipo básico de modelo cognitivo de determinación jerárquica. La idea de que las explicaciones cognitivas deben consistir en teorías de cómputo sobre representaciones simbólicas se prefiguraba también en aquella presentación de Chomsky.

En realidad, la idea de Chomsky era un producto histórico de las mismas reflexiones sobre lenguajes formales, iniciadas en el racionalismo y desarrolladas

por lógicos y matemáticos, que se encarnaron, por aquellos años cincuenta, en la fría piel de los primeros ordenadores digitales (von Neumann murió el mismo año de las jornadas del MIT, cuando aún estaba muy reciente su importante aportación a este desarrollo). Desde el comienzo mismo de la psicología cognitiva, la relación íntima entre el nuevo enfoque psicológico y la nueva inteligencia artificial quedó asegurada por el trabajo incansable y el entusiasmo de dos pioneros de ambos campos: Newell y Simon, que en la reunión del MIT describían su «Teórico lógico», que constituía la primera demostración completa de un teorema realizada por un programa de ordenador. Y, en la misma reunión, George Miller ponía de manifiesto limitaciones importantes del recién descubierto «sistema humano de procesamiento de la información», en su presentación sobre el mágico número  $7 \pm 2$ .

La reunión del MIT en 1956 no era un episodio aislado, sino una expresión del *zeitgeist* naciente, que reflejaba, sí, la sensación en muchos científicos de la conducta de que el viejo objetivismo era insuficiente e innecesariamente limitativo, pero sobre todo la convicción de que se contaba con un nuevo lenguaje con el que hablar de la mente. En aquel mismo año de 1956, se celebró en Darmouth una importante conferencia de los primeros visionarios que trabajaban en el desarrollo de sistemas de Inteligencia Artificial. La reunión, organizada por John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Shannon, y magníficamente descrita por Pamela McCorduck (1991), reflejaba las primeras aspiraciones de un proyecto, el de la IA, estrechísimamente vinculado hasta hoy con la psicología cognitiva —hasta el punto de que los límites entre una y otra son completamente «borrosos»— y de enormes consecuencias prácticas. En Darmouth, Newell y Simon defendieron un enfoque de la IA que deslindaba la noción de cómputo de la de cálculo, y defendieron una versión abierta de la IA, como versando sobre máquinas simbólicas en general, del tipo del «Teórico lógico», que ellos mismos habían desarrollado. También en aquel mismo año, 1956, Bruner, Goodnow y Austin (1956) formulaban con claridad un enfoque cognitivo del pensamiento, describiendo en *A Study of Thinking* las estrategias activas de elaboración de la información que los sujetos emplean en los procesos de formación de conceptos clásicos.

La segunda mitad de los años cincuenta constituye el momento histórico en que se define con claridad la naturaleza de la psicología cognitiva emergente. Una influencia decisiva en ese proceso de definición fue la del lingüista Noam Chomsky, que en 1957 publicaba *Syntactic Structures*, y en 1959 su demoledora crítica a la posición de Skinner sobre el lenguaje en *Verbal Behavior*. Conviene que nos detengamos brevemente en la significación de estas dos aportaciones para comprender la influencia de Chomsky.

La definición de la competencia lingüística en *Syntactic Structures* se convirtió enseguida en una especie de paradigma o modelo ejemplar de la nueva definición de la mente que entreveía la psicología emergente. Chomsky presentaba al sujeto lingüístico como una entidad, o sistema, esencialmente *formal*, definido por un conjunto de *reglas* precisas sobre *representaciones simbólicas* de carácter discreto. Definía la competencia lingüística como un *conocimiento* de reglas y representaciones, totalmente independiente del marco de la conciencia, y

con una propiedad inherente de creatividad, que permitiría en principio construir infinitas oraciones gramaticales a partir del conjunto finito, y formalmente definido, de tales reglas y representaciones no-conscientes. Además demostraba que los modelos asociacionistas y las gramáticas de estados finitos eran incapaces de dar cuenta de esa creatividad intrínseca al lenguaje. Establecía así un marco conceptual en el que iban a situarse enseguida los primeros psicólogos cognitivos: un marco que acentuaba el papel del conocimiento y de la actividad interna en el origen de la conducta, la naturaleza inherentemente creativa, jerárquica y componencial de los procesos cognitivos y la existencia de un depósito de reglas y representaciones, no directamente accesibles a la indagación consciente, y que serían requisitos necesarios para la nueva definición computacional de la mente.

Por otra parte, la crítica más bien desabrida de Chomsky a *Verbal Behavior* tuvo una relevancia histórica cierta tanto por su influencia como por su contenido: en esencia, consistió en demostrar que las propiedades *extensionales*, con las que Skinner había tratado de caracterizar la conducta verbal y el significado (propiedades tales como la «fuerza de respuesta») eran absolutamente irrelevantes para definir la naturaleza del lenguaje. Para Chomsky, el intento de hablar de la «conducta verbal» —que no del lenguaje— prescindiendo de la mente había terminado en un sinsentido. Con una notable competencia crítica, Chomsky se convertía en paladín de lo que bastantes psicólogos pensaban en voz más baja y quizá con argumentos menos definidos, al tiempo que golpeaba la posición del conductismo más radical —y quizá también más coherente— donde más podía doler el golpe: en su intento fundamental de reducir a un vocabulario puramente extensional todos los productos de la mente, entre los que el lenguaje ocupa un lugar prominente.

Así, en los años en que se constituía la nueva psicología cognitiva, Chomsky presentaba un enfoque formalista de lo mental, que se había opuesto con éxito a las pretensiones conductistas de explicar el lenguaje, y que no sólo era compatible sino modélico para el mecanicismo naciente de la nueva psicología. Del mismo modo que una máquina de Turing es una máquina formal, la gramática generativo-transformacional también lo es. El primer enfoque chomskiano situaba la *sintaxis* en el núcleo de la explicación del lenguaje. Se ofrecía así como paradigma de los que habrían de ser enseguida —especialmente en el modelo C-R— *modelos esencialmente sintácticos de la mente*. Esos modelos formalistas eran hijos del mismo *pathos* racionalista que guiaba las poderosas intuiciones de Chomsky y que había originado el viejo sueño de construir una máquina universal «capaz de resolver cualquier duda sobre los resultados del razonamiento» y de manejar los símbolos escritos en un supuesto lenguaje universal del pensamiento: la intuición histórica de racionalistas como Descartes y Leibniz, que fue a parar en los sistemas digitales de computación de propósitos generales, que eran ya una realidad a finales de los años cincuenta.

Lo que en la formulación de Chomsky puede señalarse como «gramaticismo» o «sintacticismo» tuvo un equivalente claro en el «logicismo» que fue una marca de fábrica de la nueva psicología en sus primeras fases. La influencia del modelo lingüístico de explicación se reconocía explícitamente en una obra que ya era una formulación clara del enfoque cognitivo en psicología: *Plans and the*

*Structure of Behavior* de Miller, Galanter y Pribram (1964). Probablemente no hay ninguna otra obra psicológica en que se exprese de manera más evidente el intento que caracterizaba al planteamiento naciente: el esfuerzo por insertar los viejos conceptos mentalistas de la psicología natural (tales como los planes, las intenciones y las imágenes) en los nuevos modelos explicativos de carácter cibernético o laxamente computacional. Del mismo modo que el propósito esencial del conductismo fue el de traducir el vocabulario de la mente a términos puramente extensionales, «desmentalizándolo» por así decirlo, el de la psicología cognitiva consistía en asimilar ese vocabulario a los esquemas trazados por los nuevos algoritmos de cómputo y bucles de retroalimentación.

En cierto modo, la fuerza de la nueva psicología (quizá como la de algunas formaciones políticas de amplio espectro) se basaba en su «equivocidad constitutiva», en un gran equívoco, que ha sido destacado lúcidamente por Kintsch, Miller y Polson (1984): el equívoco consistía —y consiste— en mezclar los conceptos propositivos y mentalistas de la psicología natural con las categorías del paradigma C-R, es decir, con las categorías derivadas de la noción de cómputo. Términos que sólo tienen sentido en un plano *personal* de explicación, como «plan» y «propósito», o que están inevitablemente vinculados a la fenomenología de la conciencia, como «imagen», se entremezclan e insertan en las mismas explicaciones en que se emplean nociones tales como las de TOTE (el bucle definido en *Plans and the Structure of Behavior*) o cómputo, que sólo pueden tener sentido en un nivel subpersonal de explicación.

Esa propensión equívoca se manifiesta con mucha claridad en el libro de Miller, Galanter y Pribram: los autores recurren constantemente a ilustraciones muy expresivas de sus ideas, vinculadas a las intuiciones intencionalistas que los humanos tenemos como psicólogos naturales. Nos parece que hablan de los planes de que somos conscientes, de las imágenes que nos vienen a la mente... pero su aportación fundamental consiste en definir la forma general (la organización TOTE) de los mecanismos subpersonales relevantes para estudiar los procesos de conocimiento. La convivencia de estas dos fuerzas, como la de liberales y socialistas en algunos partidos europeos, era un factor poderoso para explicar el atractivo político de los nuevos planteamientos sobre la mente, pero dejaba abierta la puerta a numerosos problemas de las explicaciones cognitivas (como los relacionados con los papeles respectivos de formas y contenidos en la dirección de los procesos mentales, con la tarea real de la conciencia en el curso del procesamiento, o con el estatuto cognitivo de las imágenes mentales) que aún hoy siguen sin estar resueltos conceptualmente del todo.

La formulación más clara del nuevo modelo computacional de explicación fue la que realizaron, en 1958, Newell, Shaw y Simon, en «*Elements of a Theory of Human Problem Solving*». En aquel artículo se ofrecía una explicación de la solución humana de problemas basada en el supuesto de que la mente es una especie de «paquete de software escrito en lenguaje simbólico de alto nivel», y la psicología la ciencia que estudia ese software. Se prefiguraba una idea que después ha sido desarrollada mucho más concienzudamente por Newell y Simon (1976): la idea de que las mentes, como los ordenadores, corresponden a sistemas que son, a un tiempo, *físicos* y *simbólicos*. Es decir, mecanismos que se rigen

racionalmente por cadenas de símbolos, las cuales corresponden, en último término, a estados físicos que funcionan con arreglo a las leyes ordinarias de causalidad física. Tales símbolos serían esencialmente intencionales, al estar situados en mundos de objetos a los que designan o interpretan, y al tiempo se manipularían con arreglo a algoritmos computacionales bien definidos.

La aportación de Newell, Shaw y Simon al impreciso campo fronterizo entre la psicología cognitiva y la inteligencia artificial presentaba tres características muy relacionadas y que fueron prototípicas de los primeros modelos cognitivos de los años sesenta: la propensión formalista, una inclinación fuertemente logicista y, por último, una formidable ambición explicativa. Las primeras explicaciones cognitivas reflejaban la huella del viejo y ambicioso proyecto racionalista de definir un «lenguaje universal del pensamiento», inmune a los contenidos y liberado de las impurezas de los factores semánticos y pragmáticos que en el pensamiento intervienen. Del mismo modo que Chomsky había tratado de definir una especie de «lenguaje universal del lenguaje» en *Syntactic Structures*, Newell, Shaw y Simon intentaron ofrecer un sistema algorítmico capaz de resolver un amplísimo conjunto de problemas: el «*General Problem Solver*».

Así, los primeros modelos explicativos de la psicología cognitiva se basaron generalmente en el supuesto de que *es posible definir mecanismos extremadamente generales de inferencia, solución de problemas y generación lingüística cuyo funcionamiento estaría determinado por la forma* (se trataba, por tanto, de modelos esencialmente sintácticos). La hipótesis de la complejidad derivacional, de la que se nutrió la ciencia recién inaugurada de la psicolingüística (que no se identificaba con los intentos reduccionistas de la psicología del lenguaje anterior), era, por ejemplo, un intento de trasladar a la psicología el enfoque gramaticalista de Chomsky en su aspecto más puro. En el campo del pensamiento y el razonamiento, los primeros estudios cognitivos se basaron en el postulado logicista implícito de que «las representaciones y/o procesos mentales humanos son isomorfos respecto del sistema de reglas formales lógico o matemático (vg.: lógica de proposiciones moderna y lógica de predicados de primer orden)» (De Vega, 1981, p. 347).

De este modo, las explicaciones sintácticas y formalistas definieron la primera fase de la psicología cognitiva. En la dialéctica que siempre ha caracterizado a este enfoque, que es la dialéctica entre sintaxis y semántica, entre reglas y contenidos, entre forma e intencionalidad, fueron los tres primeros miembros de estos pares —sintaxis, reglas y formas— los que impusieron una abrumadora presencia en los años sesenta, constituyendo una imagen del sujeto psicológico como sujeto esencialmente racional y racionalmente independiente de los contenidos procesados. Es posible que esa imagen resultara históricamente útil para corregir tanto los excesos del conductismo (en el dibujo de un sujeto reducido en gran parte a objeto de las manipulaciones y el control de su medio) como la versión irracionalista del sujeto psicológico desarrollada por muchos teóricos del psicoanálisis.

Así, el sujeto psicológico de las concepciones cognitivas de los años sesenta era una especie de pequeño monstruo racional (Rivière, 1987): un sujeto que, para hablar y comprender, emplearía sus conocimientos tácitos de las reglas mor-

fosintácticas de la gramática generativa y transformacional (Miller y McKean, 1964; Savin y Perchonock, 1965; Mehler, 1963; McMahan, 1963; Slobin, 1966; Gough, 1965, 1966), que crecería acercándose progresivamente a un modelo final de sujeto competente en los esquemas de inferencia definidos por las lógicas de clases, proposiciones y relaciones (Inhelder y Piaget, 1955); en realidad, un procesador óptimo de la información (Levine, 1975), que elaboraría conceptos limpios, semejantes a clases lógicas, mediante estrategias activas de formación y comprobación de hipótesis (Bruner, Goodnow y Austin, 1956), y realizaría operaciones lógicas sobre la información al razonar (Hunter, 1957). Un sujeto que ajustaría sus juicios de semejanza al modelo métrico euclidiano (Attneave, 1950; Shepard, 1962) y sus predicciones intuitivas a las leyes bayesianas del cálculo de probabilidades (Edwards, 1968; Peterson y Beach, 1967)... un modelo, en suma, de virtudes racionales.

Esta imagen, sólo un poco caricaturesca, expresa la realidad histórica de una psicología en la que se impusieron, en un primer momento, modelos canónicos —lógicos, matemáticos, lingüísticos— de naturaleza racional, que eran los más a mano para explicar las teorías sobre las cuales debían explicitarse algoritmos concretos (por seguir la distinción de Marr, 1982) de los nuevos modelos de procesamiento. Pero, por otra parte, refleja también nuestra caricatura otra realidad histórica importante: la de una nueva psicología que salía de su reducto particularista para vincularse a otras ciencias, y muy específicamente a ciencias formales, en un proyecto común, aún titubeante y naciente por aquellos años, de desarrollar una ciencia cognitiva unitaria. Sin duda, el enfoque cognitivo ha sido el más «importador de conceptos ajenos», y especialmente de sistemas canónicos formales, de todos los desarrollados a lo largo de la historia de la psicología. Sin embargo, a esta primera fase de importación conceptual de modelos canónicos y racionales de la mente siguió pronto otra en que, sin renunciar al modelo computacional, la psicología cognitiva tuvo que empezar a reconocer las prerrogativas de los significados —y no sólo de las formas—, los contenidos y los aspectos intencionales de lo mental. Una fase que desarrollaría un modelo mucho más flexible —pero también, en cierto sentido, más limitado— del sujeto cognitivo.

### **Rasgos de desarrollo del enfoque cognitivo en psicología: los años setenta y ochenta**

Siempre ha sido una intención osada la de tratar de historiar el presente, sobre todo cuando es el de un sistema aún en fase de rápido crecimiento. El proceso de desarrollo de la psicología cognitiva en los años setenta y ochenta, determinado por índices cuantitativos, tales como los relacionados con la ley de crecimiento exponencial de la producción científica, establecida por Price (1973), demuestran que la psicología cognitiva creció con una rapidez inusitada en los momentos inmediatamente posteriores a su definición como alternativa para estudiar objetivamente la mente (Rivière, 1987). La magnitud de ese crecimiento exponencial puede explicarse históricamente por el efecto sinérgico de la atrac-

ción de las nuevas tecnologías del conocimiento, a las que el enfoque cognitivo estuvo vinculado desde el principio, y por el hecho de que la nueva psicología se mostrara, al tiempo, más respetuosa con la complejidad que intuitivamente atribuimos a lo mental, más cercana, al menos en apariencia, a supuestos optimistas tradicionales sobre la naturaleza humana (como los de *actividad, autodirección y racionalidad*) y más objetiva que cualquiera de las alternativas mentalistas anteriores en psicología, siendo capaz de resolver aparentemente la vieja antinomia entre mente y objetividad.

De forma parecida a lo que sucede en el desarrollo físico y psicológico de los organismos, que generalmente no sólo crecen cuantitativamente sino que sufren cambios cualitativos en ese desarrollo, sin perder por ello su identidad, el nuevo enfoque sufrió variaciones importantes en los años setenta y ochenta, sin perder en ellas su identidad cognitiva y la referencia, más o menos lejana según los investigadores, a las nociones de cómputo y procesamiento de la información. Debemos limitarnos aquí a establecer los aspectos más generales y sobresalientes de esos cambios cualitativos. En una síntesis rápida señalaremos que los cambios afectaron enseguida a la imagen logicista de «sujeto racional» que ofreció en un primer momento la psicología cognitiva, y que implicaron un nuevo equilibrio en la dinámica entre forma e intencionalidad, o entre factores sintáctico-formales, por una parte, y variables semánticas y pragmáticas, por otra. Ya desde finales de los años sesenta, los aspectos semánticos (los relacionados con los *contenidos* de lo que se procesa, y no sólo con su *forma*) comenzaron a reclamar un papel en las nuevas explicaciones de la mente, aunque hay que reconocer que no es tarea fácil la de entender ese papel desde los modelos de la mente como *sistema de cómputo* (definido por algoritmos formalmente guiados) *de representaciones simbólicas*. La intrusión de estos factores tuvo mucho que ver con un tercer aspecto que definió a la nueva etapa: muchos investigadores renunciaron a la ambición primera de ofrecer modelos mecanicistas de la mente extremadamente generales —que pronto demostraron sus insuficiencias— dando lugar a una psicología de aspecto más bien «mosaico» y minimalista en muchas de sus aportaciones.

El primer cambio al que nos hemos referido implicó romper con la ecuación primera entre «sujeto algorítmico» y «sujeto hiper-racional». Los psicólogos cognitivos se dieron cuenta, enseguida, de que es posible formular explicaciones algorítmicas y modelos computacionales de conductas que no corresponden a modelos lógicos o canónicos de racionalidad. Desde finales de los años sesenta, empezó a perfilarse una imagen más versátil, flexible y limitada del sujeto cognitivo que la que se tuvo en un primer momento: se demostró que las estructuras con que produce y comprende el lenguaje están agujereadas por difusas y punzantes influencias semánticas y pragmáticas (Clark y Clark, 1977), que su competencia final en el manejo de tareas lógicas es bastante limitada, específica de tareas y dependiente de contenidos (Wason, 1966, 1968), que emplea categorías naturales difusas y de límites imprecisos (Rosch, 1978), que su razonamiento es frecuentemente alógico (Evans, 1972) y se guía por «modelos mentales» de naturaleza semántica, más que por reglas formales (Johnson-Laird, 1983), y que sus estimaciones de semejanza y juicios intuitivos se dejan influir por sesgos irrepresentables en la métrica euclidiana o el cálculo de probabilidades (Tversky, 1977;

Kahneman y Tversky, 1973). Al crecer, el sujeto de la psicología cognitiva —siguiendo, en cierto modo, un proceso contrario al que siguen los niños— se hizo cada vez menos serio, más difícil de formalizar, menos predecible y se alejó del modelo de racionalidad canónica que se le había marcado en un principio. En cierto modo, abandonó el «estadio de las operaciones formales», que había pretendido ocupar nada más nacer.

Este proceso evolutivo —que no involutivo— tuvo mucho que ver con el segundo aspecto al que hacíamos referencia: la constatación de la importante influencia de los contenidos y factores semánticos en los procesos cognitivos. Por ejemplo, la hipótesis de la complejidad derivacional, que había jugado un papel decisivo en el origen mismo de la psicolingüística cognitiva, tuvo que ser abandonada, al demostrarse que, en ciertas condiciones semánticas, oraciones complejas, desde el punto de vista transformacional, pueden ser más fáciles de comprender —o igual de fáciles— que otras simples (Wason, 1965; Greene, 1970; Johnson-Laird, 1968; Clark, 1965; Johnson, 1967). En el estudio del pensamiento, los modelos sintácticos más ambiciosos y generalistas, como el «Solucionador General de Problemas» de Newell y Simon (1972), demostraron pronto que su ámbito explicativo era mucho más reducido que lo que se pretendió en principio, y que sólo eran capaces de explicar la solución de problemas cerrados y de poca «densidad semántica», comportándose de forma aleatoria y muy poco «inteligente» en situaciones en que se aumentaba la carga semántica de los problemas. Además, las investigaciones sobre razonamiento terminaron por convertir en un lugar común la idea de que el razonamiento humano es extremadamente sensible a los contenidos, cambios de contexto y factores pragmáticos.

Si la mente es un instrumento tan versátil que modifica esencialmente sus estrategias en función de los contenidos que procesa, y de los contextos pragmáticos en que realiza su tarea, ¿hasta qué punto es posible formular modelos generales de cómputo para ella? La nueva imagen de la mente, que se dibujaba en la psicología cognitiva de los años setenta y ochenta, se correspondía estrechamente con la imagen mosaica, fragmentaria, dependiente también de contenidos y contextos, de la propia psicología cognitiva. Bien es verdad que no dejaron de formularse, en aquellos años, *modelos generales*, con la ambición de explicar una amplísima gama de actividades mentales a partir de un conjunto limitado de supuestos. Es destacable, por ejemplo, el enorme esfuerzo de John Anderson (Anderson y Bower, 1973; Anderson, 1976, 1983) en esa dirección: armado con un arsenal limitado de proposiciones y reglas de producción (de poder computacional equivalente al de la máquina de Turing) intentó, una y otra vez, la ambiciosa tarea de formular una teoría general de la mente. Todavía no existe perspectiva histórica suficiente para valorar el éxito de esa ambición, pero sí para señalar que era excepción más que regla en la psicología cognitiva de los años setenta y ochenta (si exceptuamos, desde luego, el caso especial del conexionismo, al que nos referiremos luego).

Hay quizás otros dos rasgos que deben destacarse, en la fase de consolidación del enfoque cognitivo, y que, en cierto sentido, configuran una tendencia opuesta a ese mosaicismo de los micro-modelos de mini-mentes, al que nos hemos referido. Son dos rasgos también muy relacionados entre sí: la tendencia a

emplear nociones cada vez más molares o *macroestructurales* y la creciente demanda de *validez ecológica* en las observaciones y modelos de la mente. En cuanto a lo primero, las influencias de los significados y conocimientos previos en los modelos cognitivos, hicieron cada vez más patente la necesidad de explicitar mecanismos de «arriba a abajo» que, en cierto modo, habían sido, desde los orígenes de la propia psicología cognitiva, motivos importantes para el desarrollo de la nueva concepción de la mente. La explicitación de tales mecanismos hizo conveniente recurrir a nociones de vieja raigambre epistemológica, como la de *esquema*, que se hicieron precisas, tanto en psicología cognitiva como en IA, para explicar, por ejemplo, cómo pueden procesar los sistemas cognitivos unidades de discurso. En cierto modo, las propiedades funcionales de la mente a la que estos conceptos daban respuesta —propiedades tales como la capacidad de integrar conocimientos en totalidades significativas amplias, de anticiparse a informaciones nuevas y de añadir «contenidos no-explicitos», que constituyen inferencias necesarias o convenientes para comprender lo real o el lenguaje— eran propiedades a las que, desde un principio, había tratado de hacer justicia el enfoque cognitivo. El problema que se planteaba era que el coste que había que pagar para dar cuenta de esas propiedades era, paradójicamente, una cierta inflexibilidad que siempre han manifestado las diversas versiones de la noción de esquema, y que ha intentado ser resuelta por la reinterpretación de esas nociones en los nuevos modelos cognitivos del enfoque conexionista.

El conjunto de observaciones anteriores nos permite establecer un fenómeno histórico de singular importancia en el desarrollo de la psicología cognitiva y de las tecnologías del conocimiento: mientras que, en su primera fase, los nuevos modelos de la mente fueron principalmente *importadores netos* de nociones que provenían de otras ciencias cognitivas, y en especial de las más formalistas, en la segunda, el paisaje mucho más diverso, menos monótono y racional de la mente, que se iba configurando, reclamaba un esfuerzo de formalización a esas otras ciencias, de modo que la psicología empezó a exportar conceptos basados en una versión algo más realista de la mente.

Esa versión realista resultaba útil, por ejemplo, para los investigadores de la IA: no hay que olvidar el hecho de que la propia psicología cognitiva es la principal responsable de establecer la base semántica de datos de la ciencia cognitiva como un todo. Ello es así, en tanto que es la disciplina que tiene la misión de formular versiones explicativas de las mentes reales como sistemas de conocimiento; versiones basadas en los datos obtenidos a partir de la investigación empírica sobre la conducta o, en último término, sobre la propia mente. Probablemente no es una exageración (aunque sí un peligroso ejercicio histórico de uso de un contrafactual) decir que, si la psicología cognitiva no hubiera sido sentida como necesaria por la propia lógica interna de la disciplina, hubiera sido reclamada, antes o después, por las nuevas tecnologías del conocimiento.

### **El conexionismo: un vocabulario extensional para la mente cognitiva**

Es poco frecuente que un historiador de la ciencia pueda afirmar, con un

sentimiento de certeza, la relevancia histórica de un episodio muy reciente en la disciplina que estudia. Creo, sin embargo, que pocos historiadores de la psicología negarían que la aparición del conexionismo, la publicación en 1986 de la «Biblia» conexionista *Parallel Distributed Processing* de Rumelhart, McClelland y el grupo de investigación PDP, constituye uno de esos casos de transformación revolucionaria (más que evolutiva) que se producen ocasionalmente en las ciencias.

Desde la perspectiva de análisis empleada en este examen histórico de la psicología cognitiva, el conexionismo tiene una significación especial: implica, ni más ni menos, que *la vuelta de un lenguaje extensional, pero para hablar, esta vez, de la mente* o, quizá, del punto en que la mente como sistema funcional de cómputo establece su intersección con el sistema nervioso.

¿Un vocabulario extensional para hablar de la mente?, ¿no es esa una alternativa muy poco intuitiva? ¿no se define la mente, acaso, por una propiedad esencial de intencionalidad?, ¿qué puede implicar, entonces, un lenguaje al tiempo mental y extensional? Sí, la alternativa es poco intuitiva, de forma semejante a como es poco intuitiva la alternativa física de la mecánica cuántica. Y, del mismo modo que el paisaje propio de la mecánica cuántica es el micromundo más molecular de la física, ese plano en que casi la física deja de serlo en el sentido tradicional, así también el medio propio de los modelos conexionistas es precisamente el plano más molecular en que podemos hablar de mente, el nivel de la «microcognición», por emplear un término de Clark (1989): es decir, aquel punto en que la psicología *casi* deja de ser psicología, para empezar a ser neurología teórica. Y... sin embargo, el conexionismo parece ser *aún* psicología, porque *sí* trata de la mente. No, desde luego, de la mente fenoménica definida por la milenaria psicología natural o la secular psicología filosófica o científico-introspectiva (sí es que tal adjetivo tiene algún sentido). No de esa mente, pero sí de la *mente computacional*, definida previamente en la alternativa clásica de la psicología del procesamiento de la información, y sobre todo del paradigma simbólico, a la que nos hemos referido en estas páginas.

Sin embargo, y aún versando sobre la mente, lo cierto es que los parámetros de los que hablan los conexionistas, cosas tales como los patrones de conectividad, las unidades subsimbólicas definidas por niveles de activación, las fuerzas de conexión, los umbrales de input y output de tales unidades, no son, desde luego, objetos intencionales. No son como los esquemas, las imágenes y las proposiciones mentales de los psicólogos cognitivos clásicos. Tampoco son estímulos o respuestas potencialmente exteriorizables, como lo eran las construcciones extensionales de los conductistas. Ni son nada a lo que tengamos —ni mucho menos— un «acceso fenoménico» a través de la introspección (resulta del todo insensata la idea de acceder por introspección a los niveles de activación, por ejemplo, de esas unidades que funcionan en una maravillosa actividad paralela cooperativa-competitiva, ¿nos volveríamos locos si accediéramos a esas minucias mentales!). No son, ni siquiera parámetros realistas del funcionamiento neuronal definido a un nivel puramente fisiológico. Son, en realidad, los objetos y parámetros que definirían a la mente como sistema de cómputo, en tanto que se encarna en un tipo particular de procesador: en un sistema nervioso.

¿Hasta qué punto representa este nuevo punto de vista una auténtica rup-

*tura* con los modelos tradicionales de la psicología cognitiva?, ¿en qué, por el contrario, los continúa? Es importante destacar el hecho histórico cierto de que la pregunta por las propiedades computacionales de una red neural sencillamente *no hubiera sido posible* sin la pregunta previa sobre la mente como sistema de cómputo. La metáfora, o descripción literal, de la mente computacional ha sido la que ha hecho posible un enfrentamiento nuevo al viejísimo problema histórico de la psicología de las relaciones entre una sustancia extensional, el cuerpo, y un conjunto de funciones intencionales, al que llamamos mente. Un aspecto importante —e insuficientemente destacado— de la importancia histórica del conexionismo es el hecho de que representa, en efecto, una alternativa realista, y no dualista, a ese problema secular, que a lo largo de la historia se había mostrado como intratable. Una alternativa que sólo tiene sentido cuando se sitúa el problema en el plano de lo mental «descubierto» o «inventado» (no sé qué término sería más exacto) por el enfoque cognitivo en psicología: el plano computacional. En los modelos conexionistas, sin embargo, lo que se computan no son «símbolos» —como sucedía en los clásicos— sino variables subsimbólicas de unidades moleculares de cómputo, que actúan en paralelo y con arreglo a leyes precisas.

¿Dónde quedan entonces los esquemas y conceptos, las proposiciones y, en general, los objetos intencionales de la psicología cognitiva del paradigma simbólico y de la vieja psicología natural? Desde la perspectiva molecular del conexionismo serían sencillamente propiedades emergentes, funciones resultantes de un modo de procesamiento holístico y paralelo de un sistema que tiende a alcanzar estados de equilibrio, con arreglo a modelos probabilísticos de funcionamiento, propios de ciertos sistemas complejos. Los conceptos, los esquemas, los «conocimientos» en definitiva, «almacenados» en las clásicas memorias del modelo tradicional de procesamiento, se entienden, así, como metáforas con las que nos referimos al tipo de características globales, propiamente cognitivas e intencionales, que emergen del funcionamiento global de sistemas, o subsistemas, de procesamiento, cuando éstos actúan como totalidades.

Históricamente, el modelo conexionista se enraíza en los precedentes, de los que hablábamos en el apartado segundo, de «Un cálculo lógico inmanente en la actividad nerviosa» de McCulloch y Pitts (1940) y el provisionalmente malogrado «perceptrón» de Rosenblatt (1962). Supone, por consiguiente, la continuación de un proyecto, interrumpido durante cerca de un cuarto de siglo, pero con el que nacieron desde un principio la psicología y la ciencia cognitiva: la ambiciosa pretensión de desarrollar modelos de cómputos que fueran, al tiempo, respetuosos con los datos de conducta y con los aspectos computacionalmente relevantes del funcionamiento del sistema nervioso.

Este intento se retoma con la pretensión, declarada de forma explícita y hábil, de resolver algunas de las *anomalías* principales acumuladas por el paradigma simbólico-computacional clásico en sus treinta años de existencia: la «mente serial», descrita por éste, presenta dificultades importantes para enfrentarse a condiciones que son muy naturales y frecuentes en el medio adaptativo del hombre. Situaciones, por ejemplo, en que es necesario tener en cuenta a la vez diversas restricciones simultáneas, o en que se hace preciso reconstruir o recordar informaciones a partir de partes de ellas que pueden, incluso, no ser correctas, en que

la información está parcialmente degradada (o lo está el propio funcionamiento del sistema nervioso). Además es difícil explicar cómo esa mente serial y simbólica puede ser tan diestra como lo es la mente real, para captar rápidamente los significados globales de las situaciones, y tan flexible como la mente real, si sus esquemas y conceptos son tan discretos y rígidos como lo son las representaciones que emplean los ordenadores digitales de la «familia von Neumann». Es difícil que éstos simulen completamente el estilo heurístico de procesamiento de las mentes reales. Y su funcionamiento esencialmente secuencial no parece realista como modelo del funcionamiento de un sistema nervioso que, a pesar de sus impresionantes habilidades cognitivas, es mucho más lento que los propios ordenadores en su velocidad de transmisión.

Los modelos conexionistas tratan de ofrecer una alternativa a estas cuestiones propiamente *psicológicas*, acerca de la mente, que habían surgido como dificultades de los enfoques cognitivos anteriores. Sin embargo, plantean desde un principio sus propias anomalías como modelos de lo mental: ¿cómo es posible que surjan formas regladas, estructuras netas, formalmente definibles, de su modo de funcionamiento esencialmente estocástico? ¿Cómo puede acceder un sistema conexionista a formas rigurosas, intuiciones de necesidad lógica, estructuras gramaticales? Es cierto que la solución conexionista da solución a muchas de las anomalías del modelo C-R clásico (las paradojas sobre el aprendizaje, la falta de resistencia a la degradación del sistema cognitivo, sus dificultades para explicar fenómenos como los de accesibilidad de la memoria en función del significado, etc.). Pero también es cierto que resulta difícil explicar el funcionamiento mental de «estructuras representacionales combinatorias» (es decir, con cláusulas bien definidas, partes separables, reglas sintácticas precisas, como las que se dan en el lenguaje) si no es recurriendo a los modelos clásicos. Los defensores más enérgicos del paradigma C-R, Jerry Fodor y Zenon Pylyshyn (1988) se han apresurado a destacar esta dificultad del enfoque emergente.

La psicología cognitiva vive ahora mismo un momento apasionante de choque de paradigmas, e interrelación entre ellos. Parece haber logrado dar pasos importantes en temas nucleares de la psicología: cómo definir la mente para hacerla accesible a un análisis objetivo, cómo entenderla, incluso, como función natural de una materia extensional. En sus aún pocos años de existencia, la psicología cognitiva ha acumulado una enorme cantidad de conocimiento acerca de la mente: en gran parte, se trata además de conocimientos profundos, completamente inaccesibles a la ingenua mirada de esos astutos psicólogos naturales que somos los miembros de nuestra especie. Sin embargo, su tarea pendiente es todavía ingente. No ha logrado explicar, por ejemplo, qué sentido computacional tiene esa astucia introspectiva, ni qué papel juega la propia conciencia cuando emerge por encima de los complejos cómputos que se realizan en la sala de máquinas oscura de la mente.

## REFERENCIAS

Anderson, J.R. (1976). *Language, Memory and Thought*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

- Anderson, J.R. (1983). *The Architecture of Cognition*. Cambridge, Mass.: Harvard University Press.
- Anderson, J.R. & Bower, G.H. (1973). *Human Associative Memory*. Washington D.C.: V.H. Winston.
- Attneave, F. (1959). *Applications of Information Theory to Psychology*. New York: Holt.
- Ben-David, J. & Collins, R. (1966). Social factors in the origins of a new science: The case of psychology. *American Sociological Review*, 31, pp. 451-465 (traducción española en Tortosa, F., Mayor, L. y Carpintero, H. (1990), *La psicología contemporánea desde la historiografía*. Barcelona: PPU).
- Bever, T.G., Fodor, J.A. & Garrett, M. (1968). A formal limitation of Associationism. In Dixon, R. & Horton, D.L. (Eds.), *Verbal Behavior and General Behavior Theory*. New York: Prentice-Hall.
- Boring, E.G. (1978). *Historia de la psicología experimental*. México: Tecnos.
- Bertrano, F. (1874). *Psychologie von empirischen Standpunkt*. Leipzig (ed. en Hamburgo: Meiner, 1955).
- Bruner, J., Goodnow, J.J. & Austin, G.A. (1956). *A Study of Thinking*. New York: Wiley (traducción española: *El proceso mental en el aprendizaje*, Madrid: Narcea, 1984).
- Caparrós, A. y Gabucio, F. (1986). La aparición del paradigma cognitivo: una cuestión problemática. *Revista de Historia de la Psicología*, 7 (2), 53-58.
- Clark, A. (1989). *Microcognition. Philosophy, Cognitive Science and Parallel Distributed Processing*. Cambridge, Mass.: The MIT Press. A Bradford Book.
- Clark, H.H. (1965). Some structural properties of simple active and passive sentences. *Journal of Verbal Learning and Verbal Behavior*, 4, 365-370.
- Clark, H.H. & Clark, E.V. (1977). *Psychology and Language*. New York: Harcourt Brace Jovanovich, 1977.
- Chomsky, N. (1957). *Syntactic Structures*. La Haya: Mouton (traducción española: *Estructuras Sintácticas*, Madrid: Siglo XXI, 1974).
- Chomsky, N. (1959). A review of Skinner's Verbal Behavior. *Language*, 35, 26-58.
- Chomsky, N. (1980). *Rules and Representations*. New York: Columbia University Press (traducción española: *Reglas y Representaciones*. México: Fondo de Cultura Económica, 1983).
- Dennett, D.C. (1987). *The Intentional Stance*. Boston, Mass.: MIT. Press (traducción española: *La actitud intencional*. Barcelona: Gedisa, 1991).
- Descartes, R. (1637). *Discours de la Méthode* (traducción española: *El Discurso del Método*. Madrid: Akal, 1989).
- Edwards, W. (1968). Conservatism in human information processing. In Kleinmuntz, B. (Ed.). *Formal Representations of Human Judgment*. New York: Wiley & Sons.
- Ericsson, K.A. & Simon, H.A. (1984). *Protocol Analysis: Verbal Reports as Data*. Cambridge, Mass.: The MIT Press.
- Evans, J.St.B.T. (1972). Interpretation and matching bias in a reasoning task. *Quarterly Journal of Experimental Psychology*, 24, 193-199.
- Evans, J.St.B.T. (1982). *The Psychology of Deductive Reasoning*. London: Routledge & Kegan Paul.
- Fodor, J. (1988). *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*. Cambridge, Mass.: The MIT Press. A Bradford Book.
- Fodor, J. & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3-71.
- Gardner, H. (1985). *The Mind's New Science. A History of Cognitive Revolution*. New York: Basic Books (traducción española *La nueva ciencia de la mente. Historia de la revolución cognitiva*. Buenos Aires: Paidós, 1987).
- Gough, P.B. (1965). Grammatical transformations and speed of understanding. *Journal of Verbal Learning and Verbal Behavior*, 4, 107-111.
- Gough, P.B. (1966). The verification of sentences: The effects of delay of evidence and sentence length. *Journal of Verbal Learning and Verbal Behavior*, 5, 492-496.
- Greene, J.M. (1970). The semantic function of negatives and pasives. *British Journal of Psychology*, 61, 17-22.
- Hunter, I. (1957). The solving of three-term series problem. *British Journal of Psychology*, 48, 286-298.
- Inhelder, B. & Piaget, J. (1955). *De la logique de l'enfant à la logique de l'adolescent*. Paris: Presses Universitaires de France.
- Johnson, M.G. (1967). Syntactic position and rated meaning. *Journal of Verbal Learning and Verbal Behavior*, 6, 240-246.
- Johnson-Laird, P.N. (1968). The choice of the passive voice in a communicative task'. *British Journal of Psychology*, 59, 7-15.
- Johnson-Laird, P.N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness*. Cambridge: Cambridge University Press.
- Johnson-Laird, P.N. (1988). *The Computer and the Mind: An Introduction to Cognitive Science*. Glasgow: William Collins Sons and Co. Ltd. (traducción española de Alfonso Medina: *El ordenador y la mente, Introducción a la ciencia cognitiva*. Barcelona: Paidós, 1990).
- Kahnemann, D. & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.

- Kintsch, W., Miller, J.R. & Poison, R.E. (1984). *Method and Tactics in Cognitive Science*. Hillsdale, N.J.: LEA.
- Leahey, T. (1981). The revolution never happened: information processing is behaviorism. Comunicación presentada en la 52 reunión anual de la Eastern Psychological Association. New York.
- Levine, M. (1975). *A Cognitive Theory of Learning*. Hillsdale, N.J.: Lawrence Erlbaum.
- Marr, D. (1982). *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman and Co. (traducción española: *La visión. Una investigación basada en el cálculo acerca de la representación y el procesamiento humano de la información visual*. Madrid: Alianza, 1985).
- McCorduck, P. (1979). *Machines who Think. A Personal Inquiry into the History and Prospects of Artificial Intelligence*. New York: Freeman (traducción española: *Máquinas que piensan*. Madrid: Tecnos, 1991).
- McCulloch, W. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin for Mathematical Biophysics*, 5, 115-155 (reimpreso en Boden, M. (Ed.), *The Philosophy of Artificial Intelligence*, Oxford: Oxford University Press, 1990, 22-39).
- McMahon, M. (1963). Grammatical analysis as a part of understanding a sentence. Tesis doctoral mimeografiada. Harvard, Mass.: Universidad de Harvard.
- Mehler, J. (1963). Some effects of grammatical transformations on the recall of english sentences. *Journal of Verbal Learning and Verbal Behavior*, 2, 346-351.
- Miller, G.A., Galanter, E. & Pribram, K.H. (1960). *Plans and the Structure of Behavior* (traducción española de Rodolfo Fernández González: *Planes y estructura de la conducta*. Madrid: Debate, 1983).
- Miller, G.A. & McKean, K.E. (1964). A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology*, 16, 297-308.
- Minsky, M.L. & Papert, S. (1969). *Perceptrons*. Cambridge, Mass.: The MIT Press.
- Newell, A., Shaw, J.C. & Simon, H.A. (1958). Elements of a theory of human problem solving. *Psychological Review*, 65, 151-166.
- Newell, A. & Simon, H.A. (1972). *Human Problem Solving*. Englewood Cliffs, N.J.: Prentice-Hall.
- Newell, A. & Simon, H.A. (1976). Computer science as empirical enquiry: Symbols and search. Tenth Turing Lecture. In *Communications of the Association for Computing Machinery*, 19. Association for Computing Machinery (reeditado en Boden, M. (Ed.), *The Philosophy of Artificial Intelligence*. Oxford: Oxford University Press, 1990, 105-132).
- Peterson, C.R. & Beach, L.R. (1967). Man as an intuitive statistician. *Psychological Bulletin*, 68, 29-46.
- Price, D.J. (1973). *Hacia una ciencia de la ciencia*. Barcelona: Ariel.
- Pylyshyn, Z.W. (1980). Computation and cognition: Issues in the foundations of Cognitive Science. *The Behavioral and Brain Sciences*, 3 (1), 111-132.
- Pylyshyn, Z.W. (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, Mass.: MIT Press. A Bradford Book (traducción española: *Computación y conocimiento. Hacia una fundamentación de la ciencia cognitiva*. Madrid: Debate, 1988).
- Rivière, A. (1987). *El sujeto de la psicología cognitiva*. Madrid: Alianza Editorial.
- Rosch, E. (1978). Principles of Categorization. In Rosch, E. & Lloyd, B.B. (Eds.), *Cognition and Categorization* (pp. 27-48). Hillsdale, N.J.: Lawrence Erlbaum.
- Rosenbalt, F. (1962). Principles of Neurodynamics. New York: Spartan.
- Rumelhart, D.E., McClelland, J.L. & PDP Research Group (Eds.) (1986). *Parallel Distributed Processing Explorations in the Microstructure of Cognition*. Cambridge, Mass.: Cambridge University Press.
- Shepard, R.N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika*, 27, 2.
- Slobin, D.I. (1966). Grammatical transformation and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior*, 5, 219-227.
- Turing, A.M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 2ª serie, 42, 230-265 (reimpreso en Davis, M. (Ed.), *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvability Problems and Computable Functions*. Hewlett, N.Y.: Raven Press, 1965).
- Turing, A.M. (1950). Computing machinery and intelligence. *Mind*, 59, 433-460 (reeditado en Anderson, A.R. (Ed.) *Controversia sobre mentes y máquinas*. Barcelona: Tusquets, 1984, 11-50).
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84 (4), 327-352.
- Tversky, A. y Kahneman, D. (1984). Juicio en situación de incertidumbre: Heurísticos y sesgos. En M. Carretero y García Madruga, J.A. (Eds.), *Lecturas de Psicología del Pensamiento* (pp. 169-181). Madrid: Alianza Psicología.
- de Vega, M. (1981). Una exploración de los metapostulados de la psicología contemporánea. El logicismo. *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* (pp. 63-81). Madrid: Pirámide.
- Wason, P.C. (1965). The contexts of plausible denial, *Journal of Verbal Learning and Verbal Behavior*, 4, 7-11.
- Wason, P.C. (1966). Reasoning. In B.M. Foss: *New Horizons in Psychology*. Hordhansworth, Middlesex: Penguin.
- Wason, P.C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273-281.
- Watson, J.B. (1913). Psychology as the behaviorist views it. *Psychological Review*, 20, 158-177.
- Yela, M. (1974). *La Estructura de la Conducta. Estímulo, Situación y Conciencia*. Real Academia de Ciencias Morales y Políticas.

