

ESTATUTO EPISTEMOLOGICO DE LOS CONCEPTOS MENTALES

Por José Hierro S. Pescador

1. Antecedentes

De acuerdo con una doctrina ampliamente extendida y que usualmente se denomina «cartesiana», el carácter propio de lo mental es la incorregibilidad (1). Según esto, los conceptos mentales no serían susceptibles de disconfirmación empírica cuando son utilizados por un sujeto para hablar acerca de sí mismo y de sus experiencias. Es decir, no sería posible, por principio, que alguien se equivocara al expresar sus estados mentales o informar sobre ellos. Los informes sobre la experiencia inmediata propia diferirían de todos los demás en que poseen una «lógica privada», en el sentido de que si una persona conoce el lenguaje y es sincera, entonces, por lógica, tiene que tener la experiencia inmediata sobre la que está informado. No hay aquí cabida lógica para la equivocación o el error. Uno puede equivocarse acerca de algo externo, público, intersubjetivo y observable, pero no



JOSE HIERRO S. PESCADOR es Profesor Agregado de Lógica en la Universidad Complutense de Madrid. Ha publicado entre otras obras, *Problemas del análisis del lenguaje moral* (Tecnos, 1970), y *La teoría de las ideas innatas en Chomsky* (Labor, 1976).

(1) Véase P. Bieri, «'Philosophische Psychologie'. Ueberlegungen zu einer Begriffsbildung», p. 71; y R. Rorty; «Incorrigibility as the Mark of the Mental», citado por Bieri.

* BAJO la rúbrica de «Ensayo» el Boletín Informativo de la Fundación Juan March publica cada mes una colaboración original y exclusiva de un especialista sobre un aspecto de un tema general. Anteriormente fueron objeto de estos ensayos temas relativos a la Ciencia, el Lenguaje, el Arte, la Historia, la Prensa y la Biología. El tema desarrollado actualmente es la Psicología.

En Boletines anteriores se han publicado: *Lo físico y lo mental*, por José Luis Pinillos, Catedrático de Psicología de la Universidad Complutense; *Piaget y la psicología cognitiva*, por Juan A. Delval, Profesor de Psicología Evolutiva de la Universidad Complutense; *Modelo judicativo de la conducta*, por

acerca de sus propias experiencias o estados mentales, que, por definición, serían privados, subjetivos y no susceptibles de observación externa.

Por lo que respecta a la incorregibilidad (lo demás lo consideraré más adelante), esta posición no me parece correcta. Una deformación profesional propia del filósofo y del teórico ha conducido aquí a muchos autores a mantener una tesis que patentemente no corresponde a la realidad. Cualquiera sabe que, si no con todos, al menos con muchos de nuestros estados mentales podemos equivocarnos. ¿Cuántas veces creemos que amamos a una persona, para descubrir luego que nuestro sentimiento era otro, piedad, reconocimiento acaso...? O bien, recordamos perfectamente dónde estaba la casa con la chimenea de latón, y cuando vamos a buscarla no la encontramos. O decimos: «Sé muy bien cómo sacar la raíz cuadrada de ese número»; pero lo intentamos y no nos sale. O después de escuchar al amigo atribulado sentenciamos: «Te comprendo perfectamente»; y a continuación le damos un consejo que le sume en la perplejidad. O bien sentimos que nos duele el tobillo que nos torcimos ayer... hasta que descubrimos que el tobillo torcido era el otro, y que lo que sentimos no es dolor sino la sensación normal del juego de la articulación.

No, los conceptos mentales no son incorregibles: es indudable que uno puede equivocarse sobre sus propias experiencias inmediatas (2). ¿De dónde procede la idea de la

(2) En este sentido, por ejemplo, Armstrong, *A Materialist Theory of the Mind*, p. 100-115; Smart, *Philosophy and Scientific Realism*, p. 99 ss.; Nat-soulas, «Concerning Introspective 'Knowledge'»; y muy especialmente Nisbett y Wilson, quienes en «Telling more than we Can Know: Verbal Reports on Mental Processes», han mostrado que, para los procesos cognitivos superiores, los sujetos suelen carecer de pruebas introspectivas suficientes, las cuales sustituyen por prejuicios acerca de las supuestas causas de sus respuestas. Y esto por no mencionar todo lo que desde un punto de vista psicoanalítico podría decirse al respecto.

▷ Carlos Castilla del Pino, Profesor de Psiquiatría en la Facultad de Medicina de Córdoba; *Tareas actuales de la Psicolingüística*, por Victor Sánchez de Zavala, Profesor de Psicología del Pensamiento y el Lenguaje de la Universidad Complutense; *Posibilidades y límites de los tests de inteligencia*, por J. A. Forteza, Profesor Agregado de Psicología Diferencial de la Universidad Complutense; *Herencia y ambiente en la Psicología contemporánea*, por Mariano Yela, Catedrático de Psicología General de la Universidad Complutense; *La Psicología soviética en contradistinción con la Psicología norteamericana*, por J. L. Fernández Trespalacios, Catedrático de Psicología General de la Universidad a Distancia; *Terapia y modificación de conducta*, por Vicente Pelechano, Catedrático de Psicología Evolutiva y Diferencial de la Universidad de Valencia; *Psicología y bilingüismo*, por Miguel Siguán, Catedrático de Psicología de la Universidad de Barcelona; y *Enfermedad mental y sociedad*, por Florencio Jiménez Burillo, Profesor de Psicología Social de la Universidad Complutense.

incoregibilidad? Sin duda, del hecho de que el sujeto tiene, aparentemente, un acceso directo, inmediato y privilegiado a sus propias experiencias internas. Mis recuerdos, mis pensamientos, mis deseos, mis dolores, mis nostalgias, son mías en un sentido en el que no lo son mis brazos, mis dientes, mis ropas, mis libros o mis mujeres. Lo primero es tan mío que, si no quiero, nadie más puede tener acceso a ello; de aquí, al parecer, puede deducirse que no es posible que yo me equivoque al respecto. Ahora bien, esta deducción no es válida. Por muy directo e inmediato que sea mi acceso a mis estados mentales, nada asegura que no pueda equivocarme respecto a ellos. Esto habría que probarlo independientemente; y a mi juicio, no se ha probado.

Quedémonos ahora con el característico acceso inmediato y directo (esto es, sin intermedio de los sentidos externos) y denominemos a esto «subjetividad». Los conceptos mentales —digamos— son subjetivos. La cuestión es: ¿cómo puede construirse una ciencia como la psicología empírica a base de conceptos subjetivos, de conceptos que no se refieren a nada intersubjetivo y público, sino solamente al mundo interior y privado del sujeto? Naturalmente, asumimos que la psicología, si es ciencia, lo es en el sentido en que lo es la biología, y no en ningún otro. ¿O no lo es?

En los felices años treinta del optimismo neopositivista, Carnap propuso un programa para hacer de la psicología una ciencia empírica análoga a las demás. Igual que cualquier otra disciplina aspirante a ciencia, la psicología habría de tomar la física como ejemplo y constituirse a su modo y manera, cumpliendo así el programa fiscalista que la unificación de la ciencia requería a juicio de los miembros del Círculo de Viena. Para cumplir este programa, Carnap ponía como condición que las afirmaciones psicológicas fueran traducibles a un lenguaje en el que los términos no lógicos tuvieran como referencia estados corporales, los cuales eran definidos en términos de disposiciones para reaccionar de cierta manera a estímulos determinados (3). Para estos efectos, una afirmación psicológica sería traducible a una afirmación física siempre que aquélla pudiera ser deducida de ésta y viceversa. El programa de Carnap iba, en líneas generales y por lo que se refiere a su carácter epistemológico, de acuerdo con las

(3) «Psychology in Physical Language», en la recopilación de Ayer, *Logical Positivism*, p. 165 ss.

tendencias conductistas. De hecho, Carnap abandonó posteriormente su concepción de los términos mentales como designativos de estados corporales caracterizados por disposiciones para la conducta, pasando a considerarlos como términos teóricos o constructos hipotéticos dentro de la psicología. Según esto, tales términos se introducirían como primitivos en una teoría psicológica, y se conectarían con los términos observacionales de esa teoría por medio de reglas de correspondencia, quedando asegurada la relación entre la teoría y la observación de la misma manera que en cualquier ciencia (4) (suponiendo que la idealizada visión de Carnap sobre la estructura de la ciencia sea correcta: las investigaciones de Kuhn y otros la han desacreditado extraordinariamente). Esta nueva formulación no afecta, claro es, al punto de vista fisicalista, que sobrevive igualmente; lo único que cambia es la manera de interpretar los conceptos mentales. Por ello, la opinión de Carnap coincide todavía con lo que se ha considerado como el postulado epistemológico fundamental del conductismo: que existe una conexión lógica, y por tanto necesaria, entre los conceptos mentales y los conceptos que describen el comportamiento (5). Los primeros serían conceptos teóricos y los últimos serían conceptos observacionales. De aquí que el mentalismo formule a veces la crítica al conductismo afirmando que es posible dar una explicación psicológica sin que haya conexión lógica entre ambas clases de conceptos (6). La explicación psicológica, para los mentalistas, sería así: requeriría un nivel teórico, integrado por conceptos mentales, pero no podría vincular lógicamente este nivel con el nivel observacional. En esta medida, no se trataría de una explicación científica, en el sentido de la metodología neopositivista.

La concepción clásica del conductismo ha venido a recibir el nombre de «conductismo metodológico», a diferencia del llamado «conductismo radical». Con este sentido, la distinción está en el propio Skinner (7). El conductismo metodológico tendería a eliminar el ámbito mental en cuanto objeto de estudio, prescindiendo también totalmente de la introspección, del autoconocimiento y del mundo de lo subjetivo. En cambio, el conductismo radical no niega ni la posibilidad ni la utilidad del auto-

(4) «The Methodological Character of Theoretical Concepts».

(5) Véase A. Riviére, «El análisis experimental de la conducta y el conductismo radical como filosofía», p. 115.

(6) J. Fodor, *Psychological Explanation*, p. 13 y 88.

(7) *About Behaviorism*, p. 13 ss.

conocimiento; simplemente se hace cuestión del contenido de este autoconocimiento. Y lo que afirma es que el contenido del autoconocimiento no es la mente, no es la conciencia, no son los supuestos estados mentales, sino simplemente el propio cuerpo del sujeto que se autoobserva. De este modo, los conceptos mentales resultan tan vacíos de por sí como en el conductismo metodológico más antiguo, pues solamente son utilizables en cuanto puedan ser traducidos a conceptos de comportamiento. La novedad es que ahora se reconoce un mundo interior y subjetivo: el mundo de lo que hay bajo la piel. No deja de resultar irónico que este simple reconocimiento le parezca a Skinner suficiente para presentar el conductismo radical como una posición de equilibrio entre el mentalismo y el conductismo metodológico. Pero la verdad es que no lo es; ciertamente constituye una posición mucho más extremada que el conductismo metodológico. Al menos en este último los conceptos mentales poseen una función teórica, aunque la autoobservación quede eliminada como método de confirmación. Pero en el conductismo radical ya no cumplen ni siquiera esa función, y el hecho de conservar una introspección limitada a estados orgánicos internos no le acerca más al mentalismo sino todo lo contrario.

Se ha conectado en ocasiones con el conductismo la conocida crítica de Ryle, inspirada en Wittgenstein, contra los conceptos mentales. A decir verdad, y esto es algo usualmente malentendido, Ryle nunca negó en su obra que tenga sentido utilizar conceptos mentales, y menos aún que existan procesos mentales (8). Lo único que hizo fue criticar una cierta manera de hablar de la mente y de sus estados y procesos, a saber, aquella manera en la cual el ámbito de lo mental se considera como un mundo paralelo a y comparable con el mundo de lo físico o material. Contra esta consideración de la mente, que Ryle denominó «mito del espíritu en la máquina», y cuyo origen principal atribuía a Descartes, Ryle desarrolló una laboriosa y prolija argumentación, cuyo fundamental propósito era mostrar que tal enfoque constituye un error conceptual del que se derivan numerosas y variadas falacias y dificultades de todo tipo. El error consistiría en categorizar de forma equivocada todo lo relativo a la mente. Por analogía con nuestra categorización de lo corporal, lo mental vendría categorizado en los mismos términos: la mente se-

(8) *The Concept of Mind*, p. 23 de la edición de Penguin.

ría una cosa distinta del cuerpo, pero una cosa al cabo; constituiría un todo organizado como lo es el organismo; habría procesos mentales como hay procesos orgánicos, etc.

A decir verdad, la consideración dualista del hombre como compuesto de cuerpo y espíritu no se debe particularmente a Descartes ni siquiera en la filosofía; su origen es mucho más remoto, y antes que una doctrina filosófica sobre el hombre constituye un mito básico de la cultura universal y aparece en todos los niveles del discurso. Por esta razón, la crítica de Ryle quedaba reducida en su eficacia según los límites de la posición filosófica contra la que iba dirigida, límites que parecen trazados un tanto *ad hoc*. De hecho, Ryle no se metía para nada en psicología ni abordaba la cuestión de la función científica de los conceptos mentales. Se limitaba a establecer la lógica peculiar, esto es, los caracteres semánticos generales, de una porción determinada del lenguaje ordinario, aquella que se refiere a los fenómenos mentales. Intentando evitar la reducción de lo mental a lo físico, Ryle presenta lo primero como un conjunto de características de diverso tipo que presenta el comportamiento. Esta posición, que se denomina a veces «conductismo lógico», no es, por tanto, una doctrina de filosofía de la ciencia acerca de la psicología como son el conductismo metodológico y el conductismo radical, sino que constituye una tesis filosófica sobre el análisis semántico (o lógico, en un sentido amplio de este término) de una porción del lenguaje ordinario.

Una tendencia más reciente a favor de eliminar los conceptos mentales es la que identifica los estados mentales con estados neurofisiológicos, especialmente cerebrales. Quienes defienden esta posición lo hacen, muchas veces, para superar las dificultades del conductismo, esto es, para salvar aquellos casos en los que no hay comportamientos claros a los que reducir los conceptos mentales (9). De esta manera, la reducción al nivel neurofisiológico hace de lo mental un campo de investigación científica, aún subdesarrollado y poblado de incógnitas, pero metodológicamente tan justificable como el estudio de cualquier porción del organismo humano. El precio que se paga, como en el conductismo radical, es la pérdida de los fenómenos mentales en cuanto ámbito peculiar de la vida hu-

(9) Así, Smart, en *Philosophy and Scientific Realism*, p. 88 ss. También defiende esta posición Armstrong, *A Materialist Theory of the Mind*, y «The Causal Theory of the Mind».

mana. Se está concediendo, de hecho, que la psicología científica no puede tener como objeto ese orden de fenómenos: que, o bien ha de limitarse al estudio de la conducta en sus condicionamientos ambientales y genéticos, o bien ha de dejar paso a la neurofisiología.

Es lo que acontece en Quine. Para él, las entidades mentales sólo son aceptables «si se conciben como mecanismos físicos hipotéticos y si se asumen con el ojo firmemente puesto en el blanco de la sistematización de los fenómenos físicos» (10). Y añade: «Debemos asumirlas con la esperanza de someterlas un día a una explicación física». La voluntad reduccionista es patente. No se trata de que las entidades mentales puedan coincidir con estados neurofisiológicos; se trata de que no pueden ser concebidas de otra manera. Pero a falta de información sobre tales estados, el vocabulario mentalista sigue conservando una función estimulante (11). Su principal justificación, o acaso la única, sería su familiaridad, su profundo arraigo en el discurso cotidiano. Pero ello no hace más claros a los conceptos mentales ni otorga más seguridad a la introspección. Más bien, concibiéndolos como conceptos neurofisiológicos, evitamos enojosos problemas como el de si la imagen mental de un triángulo es equilátera o no; si la imagen es un cierto estado neural, entonces no tiene sentido preguntarse cómo son sus ángulos: por definición, un estado neural no tiene ángulos. Así, Quine puede prescindir de conceptos como el de idea (12) y ofrecer como programa una epistemología naturalizada, es decir, sujeta a las exigencias de la ciencia natural, para la cual las pruebas en favor de las afirmaciones psicológicas habrán de buscarse necesariamente en la experiencia sensible, intersubjetiva y comunicable (13). Acaso el resultado de repudiar así las entidades mentales sea, como él había pretendido en otro lugar, eliminar las barreras entre lo público y lo privado de tal modo que sólo resten «grados variables de privacidad en los acontecimientos del mundo físico» (14). En conformidad con esto, la conciencia podría considerarse como la facultad de responder a las respues-

(10) «Reflexiones filosóficas sobre el aprendizaje del lenguaje», p. 15; *The Roots of Reference*, p. 33-34.

(11) *The Roots of Reference*, p. 33.

(12) Se encontrará un resumen más detallado de la posición de Quine en mi libro *La teoría de las ideas innatas en Chomsky*, p. 106 ss.

(13) «Epistemology Naturalized».

(14) «On Mental Entities», p. 123.

tas propias (15), y esta facultad podría presumiblemente descomponerse en un conjunto de disposiciones, las cuales, según el programa de Quine, no serían otra cosa que estados neurofisiológicos (posiblemente no conocidos todavía).

2. Tesis

He hablado al comienzo del acceso inmediato y directo que el sujeto parece tener a sus estados mentales y que hace de los conceptos mentales conceptos subjetivos. A mi juicio, ha sido uno de los grandes errores de la filosofía de la mente el haber considerado este acceso como un fenómeno cognoscitivo, al haber pensado que la inmediatez de los estados mentales es una inmediatez de conocimiento y que lo mental constituye un ámbito privilegiado para éste. Los abusos de la introspección y de la psicología fenomenológica proceden de ahí. La inmediatez, sin embargo, es más bien una inmediatez de acción, de experiencia: los estados mentales nos son inmediatos porque somos sujetos de ellos. Soy yo el que amo, pienso, recuerdo, decido, siento dolor o nostalgia, etc. Son mis estados mentales porque soy yo el que los experimento. Por esta razón, los conceptos mentales pueden considerarse como *conceptos acerca de la estructura del ser humano en cuanto sujeto de experiencias*. De aquí que la conciencia, concepto básico entre los mentales, pueda construirse como la noción básica de sujeto de experiencias (16). Así entendidos, los conceptos mentales no son reducibles ni a conceptos de conducta ni a conceptos neurofisiológicos. No son reducibles a conceptos de conducta porque en la conducta, vista desde fuera, no se nos da un sujeto de experiencia sino un objeto determinado, un organismo de cierto tipo, que interacciona con su medio de cierta manera. Pero los conceptos mentales no surgen espontáneamente en una descripción (externa) de la conducta sino en el intento de su comprensión, que requiere verla haciéndose por el propio sujeto agente. Tampoco son reducibles los conceptos mentales a conceptos neurofisiológicos, porque estos últimos se refieren a estados físicos del sistema nervioso, y aun cuando a todo estado mental corresponda necesariamente un

(15) *Ibidem*.

(16) En este sentido decía Ortega metafóricamente que «todo, mirado desde dentro de sí mismo, es yo», en su prólogo a *El Pasajero*, tomo VI de las *Obras Completas*, p. 252.

determinado estado neural, esos estados no son estados del sujeto de experiencia en cuanto tal, sino solamente estados de una porción del organismo que es ese sujeto de experiencia. Considérese, por ejemplo, el concepto de imagen que se ha mencionado a propósito de Quine. Una imagen de triángulo es un cierto estado neural. Pero ¿por qué este estado neural lo experimento como una imagen de triángulo? Porque siento algo parecido a lo que siento cuando veo un triángulo. El concepto de imagen se refiere a mi experiencia en cuanto tenida por mí. El concepto de estado neural se refiere, en cambio, a un cierto estado de cosas en cierta porción del mundo físico, a saber, en mi cerebro. Se trata de cosas totalmente distintas.

Lo anterior no significa que los conceptos mentales designen objetos peculiares en un mundo paralelo al mundo físico. En este sentido, la crítica de Ryle es correcta, y construir el vocabulario mental bajo el modelo de la designación es algo que ya quedó suficientemente desacreditado por Wittgenstein (17).

De un lado, pues, los conceptos mentales poseen una función que consiste en suministrar una comprensión de la conducta conectándola con la experiencia que el propio sujeto tiene en cuanto agente de aquélla. Los conceptos mentales describen aspectos de esta experiencia, y en principio son, por tanto, puramente subjetivos. Nada significa que el vocabulario mental lo hayamos aprendido en conexión con la conducta, pues ello no quiere decir que tal vocabulario designe aspectos del comportamiento. Lo hemos aprendido en conexión con la conducta porque el lenguaje se aprende intersubjetivamente, y sólo la conducta suministra un criterio de corrección para el uso del lenguaje. Cuando el niño aprende a decir «me duele» es ciertamente en conexión con alguna manifestación externa de dolor. Pero esta expresión no describe ninguna manifestación ni conducta del niño, sino la cualidad específica de una cierta experiencia suya en cuanto vivida por él (cualidad, por cierto, acerca de la cual puede equivocarse, en contra de lo que parecen pensar Wittgenstein y los defensores de la incorregibilidad de los conceptos mentales) (18).

Hasta aquí, una primera función de los conceptos mentales: expresar y describir (a menudo ambas cosas al tiem-

(17) Cfr. *Philosophische Untersuchungen*, secciones 244 ss. y 292 ss.

(18) Cfr. la obra citada en la nota anterior, sección 246.

po) la experiencia del sujeto tal y como éste la vive, y en esta medida suministrar una comprensión de su comportamiento más allá de lo externamente visible. La conducta es aquí el punto de partida, y los conceptos mentales constituyen instrumentos interpretativos para su más profunda comprensión. Hay que recordar aquí el interés que vuelve a despertar en las ciencias humanas el concepto de comprensión (*Verstehen*), y aunque personalmente estoy convencido de que su rendimiento en la explicación científica es nulo, pienso, sin embargo, que es una de las categorías filosóficas básicas, a pesar de los claros peligros que comporta especialmente en su aplicación a la conducta ajena, peligros que adecuadamente ilustra su historia desde Dilthey a Gadamer pasando por Husserl y Heidegger, e incluso por el propio Wittgenstein (19).

La cuestión surge cuando intentamos hacer ciencia sobre el comportamiento. ¿Vamos a aceptar de alguna manera la remisión a ese nivel subjetivo? Una de las aspiraciones históricas de la psicología ha sido justamente ésta: intentar objetivar lo subjetivo. Mas se trata de una aspiración incompatible con el carácter científico de la psicología. Aquí sí hay que reconocer que la conducta no sólo es punto de partida sino también obligado punto de llegada. Pues las pruebas de cuanto en ciencia se diga han de ser intersubjetivas y, por tanto, comprobables. Naturalmente, hay que tomar esto sin exceso de ingenuidad. A veces no está muy claro lo que prueban las pruebas, y en otras ocasiones la teoría ya predetermina qué tipo de pruebas son aceptables y cuáles no, como la más reciente filosofía de la ciencia se ha encargado de señalar (20). Pero al cabo, pruebas ha de haber, y éstas han de ser comunicables y accesibles a todos, por lo que, para el caso que nos ocupa, parece que sólo podrán encontrarse en el comportamiento. ¿Conservaremos entonces los conceptos mentales? Si estamos dispuestos a conservar como objeto de investigación psicológica la conducta humana sin prescindir de lo que tiene de humana, y por consiguiente en su peculiaridad, parece inevitable conservar los conceptos mentales. Prescindir de ellos a la manera del conductismo radical equivale, primero, a considerar la conducta huma-

(19) Véase T. McCarthy, «The Operation Called *Verstehen*: Towards a Re-definition of the Problem».

(20) Cfr. Kuhn, *The Structure of Scientific Revolutions*; Feyerabend, «Explanation, Reduction and Empiricism»; Hanson, *Patterns of Discovery*.

na como otro orden más de acontecimientos físicos, negándole la peculiar cualidad que posee para el sujeto en cuanto experiencia vivida. Pero además, ocurre que la explicación conductista no es viable ni siquiera para el comportamiento animal en cuanto se admita que los animales son autómatas, esto es, que su conducta obedece a las relaciones matemáticas que caracterizan al tipo de objeto matemático que se denomina «autómata». Aunque esta hipótesis parece chocante, resulta extremadamente plausible en cuanto se piensa que los ordenadores digitales son autómatas (o si se prefiere: realizaciones de autómatas) y que pueden señalarse numerosas formas de conducta que los animales tienen en común con ellos; por ejemplo: la respuesta a estímulos, el recordar, el reflejo condicionado, el aprendizaje de juegos, la deducción de teoremas (por lo que respecta al animal humano), etc. Pues bien, Nelson ha mostrado que, admitido esto, el conductismo, como forma de explicación de la conducta animal, es falso (21).

Téngase en cuenta que al afirmar que todo animal es un autómata lo único que se quiere decir es que ser un autómata es condición necesaria para ser un animal, no que ser un animal equivalga a ser un autómata. Es decir, no se define «animal»; simplemente se da una condición necesaria para ser animal. Tampoco se aceptan implicaciones deterministas o mecanicistas, pues un autómata puede no ser determinista. Y sobre todo, no hay que confundir un concepto formal como es el de autómata con una realización material de un autómata como es un ordenador o un animal. No se trata, en consecuencia, de equiparar al hombre con la máquina ni de explicar la mente como un ordenador. No es el hombre el que se parece al ordenador, sino éste el que se asemeja a aquél. Y la semejanza no está en que ambos sean materiales; la semejanza se encuentra en que ambos procesan la información, transforman las entradas en salidas, de acuerdo con una determinada estructura interna que, en principio y al margen de su realización material (electrónica en el caso de un ordenador, neurofisiológica en el caso de un hombre), es una estructura puramente formal. Por eso afirma Nelson que una teoría de la mente como autómata no sería ni dualista ni materialista, sino simplemente naturalista (22).

¿Qué lugar ocupará en la ciencia psicológica esta es-

(21) «Behaviorism is False».

(22) Artículo citado, p. 451.

estructura interna? Dado que no es accesible a la observación directa, parece que pertenecerá al nivel teórico de la psicología. Los conceptos mentales habrán de construirse como conceptos teóricos pertenecientes al *explanans*, la explicación, mientras que la conducta constituirá lo que hay que explicar, el *explanandum*, por tanto, la prueba empírica de lo que se afirme en la teoría a la que recurra el *explanans*. Los conceptos mentales estarán contenidos en éste como conceptos teóricos, y constituirán los elementos de esa estructura interna que puede interpretarse, de una posible manera, como un autómeta. Que esta estructura posea una determinada realización neurofisiológica es tal vez necesario, pero no las confunde. Las características de la organización neurofisiológica no son características de la organización mental, ni viceversa. Cuando el sujeto ama, piensa o imagina, tiene una determinada experiencia que no es de ningún modo identificable con la que tendría si pudiera observar los correspondientes estados neurales de su cerebro. Con esto, tal vez no consigamos una epistemología naturalizada en el sentido de Quine, esto es, una epistemología incluida en la ciencia natural, pero sí obtenemos una epistemología «formalizada», es decir, incluida en las ciencias formales, exactamente en la teoría de autómetas. Así, tener una mente será simplemente poseer un determinado tipo y grado de organización (cualquiera que sea la realización material de esta organización) (23).

Ahora bien, esto no puede hacernos olvidar que ese lenguaje teórico que expresa esa organización puede ser utilizado, *fuera de la ciencia psicológica*, como un lenguaje que exprese y describa (a veces ambas cosas al tiempo) la experiencia que el sujeto tiene de sus propios estados mentales. De acuerdo en que esta autodescripción introspectiva puede ser equivocada, y en que raramente es lo bastante inmediata como para ser fiable. Las críticas aquí hay que dirigirlas tanto contra la introspección clásica como contra la experiencia interna de la psicología fenomenológica, como ha mostrado Pinillos (24). El problema fundamental consiste en la dificultad de objetivar algo que, además de ser subjetivo, está en perpetuo cambio y tránsito, la conciencia. Desde el momento en que se intenta describir sus contenidos, desaparece la conciencia, que es acción, sujeto, y queda un objeto integrado por re-

(23) Nelson, lugar citado en la nota anterior.

(24) «La psicología fenomenológica», p. 498.

cuerdos e interpretaciones frecuentemente pasados a través del lenguaje. Pero esto se refiere, claro es, a la utilización científica de esa autoexperiencia. Aun con toda la carga interpretativa adquirida desde fuera, el lenguaje mentalista es utilizado por el sujeto para expresar y describir sus experiencias y, por tanto, para comunicarse sobre ellas con los demás, así como para orientarse él mismo acerca de sí mismo por medio de una visión general de su propia estructura como persona, esto es, como sujeto de experiencia. Los conceptos mentales, que en la psicología científica constituyen el lenguaje teórico necesariamente conectado al lenguaje de la conducta, en la autocomprensión sirven para expresar y describir la impresión, más o menos inmediata según los casos, que la persona tiene de sí misma en cuanto sujeto de experiencia. Sellars ha imaginado las condiciones en las que un lenguaje mentalista pasaría de tener un uso teórico en la explicación de la conducta ajena a adquirir un uso informativo sobre la propia experiencia del que habla (25). Aunque habría que investigar si filogenéticamente ha ocurrido así, en todo caso parece claro que actualmente la ontogénesis de los conceptos mentales procede de esta manera. Aprendemos esos conceptos en el contexto de conductas típicas, pero al utilizarlos en primera persona les damos sin querer una referencia inmediata a nuestra propia experiencia como sujetos y al margen de nuestro comportamiento. Por ejemplo, cuando el niño aprende a utilizar aquellas expresiones en las que entra el verbo «doler» o el término «dolor», lo aprende en conexión con formas de conducta como los quejidos, el lloro, la caricia del lugar dolorido, etc., comportamientos que advierte tanto en otros como en sí mismo. Pero pronto se da cuenta de que el referente de «dolor» no es la conducta que lo expresa, sino lo que él siente, y que no puede, en cambio, advertir en los demás excepto a través del comportamiento ajeno. En cuanto empiece a entender frases como «No llores aunque te duela», se dará cuenta de que hay un dualismo ineludible entre lo que él siente, que en este caso es dolor, y las maneras de manifestarlo. El aprendizaje de los conceptos mentales es intersubjetivo; pero esto no implica que los conceptos mentales no se utilicen para describir algo propio y privado del sujeto y, en este sentido, que no tiene nada de espacial, algo interno. Wittgenstein, quien subra-

(25) «Empiricism and the Philosophy of Mind», p. 188-189.

yó adecuadamente lo primero, parece a veces mantener erróneamente lo segundo, es decir, que los conceptos mentales no pueden utilizarse con esa función descriptiva, sino todo lo más con una función expresiva que convertiría, por ejemplo, a las expresiones verbales de dolor en manifestaciones de dolor como el llanto (26). Podríamos decir, acaso, que el sentido de los conceptos mentales procede de las condiciones (intersubjetivas) de su aprendizaje, mientras que su referencia sólo aparece cuando son utilizados por cada sujeto en la propia autodescripción. Es justamente esta función que los conceptos mentales cumplen en la autodescripción del sujeto lo que hace a esos conceptos más problemáticos, pues el filósofo, el psicólogo teórico, o simplemente el aficionado a pensador, al punto se preguntará qué pasa con ese mundo de estados y procesos mentales, en qué consiste esa estructura formal que es la mente y cuáles son las leyes que rigen su funcionamiento. Con ello, lo que en principio era un recurso explicativo (*explanans*) para la conducta (*explanandum*) se convierte en algo que a su vez hay que explicar (un nuevo *explanandum*), y para lo cual hay que encontrar ahora una nueva explicación (otro *explanans*). Y aquí estamos ya en pleno desvarío mentalista. Nos sentiremos tentados de intentar explicar el mundo de lo mental, que no es intersubjetivo, para lo que buscaremos una explicación que *a fortiori* tampoco será susceptible de comprobación intersubjetiva. Tal es lo que a mi juicio acontece con la teoría chomskiana de las ideas innatas.

En resumen. Pienso que la peculiaridad epistemológica de los conceptos mentales se halla en que estos cumplen a la vez una función mediata como conceptos teóricos en la explicación científico-psicológica de la conducta, y una función inmediata en la expresión y autodescripción del individuo como sujeto de experiencia. Y que el peligro y la dificultad de estos conceptos proviene de confundir ambas funciones. La explicación de la conducta, que es el objeto de la psicología científica, es una cosa; la autocomprensión del individuo como sujeto de experiencia, que es objeto de la reflexión filosófica, otra muy distinta. Aunque en ambas intervengan los conceptos mentales.

(26) *Blue and Brown Books*, p. 67; «Wittgenstein's Notes for Lectures on 'Private Experience' and 'Sense Data'», p. 301-2; *Philosophische Untersuchungen*. secc. 244.

BIBLIOGRAFIA

- Armstrong, D. M.: *A Materialist Theory of the Mind*, Routledge, Londres, 1968.
- «The Causal Theory of the Mind», *Neue Hefte für Philosophie*, 11, 1977.
- Bieri, P.: «'Philosophische Psychologie'. Ueberlegungen zu einer Begriffsbildung», *Neue Hefte für Philosophie*, 11, 1977.
- Carnap, R.: «Psychologie in physikalischer Sprache», *Erkenntnis*, 1932-33 (trad. inglesa en Ayer, *Logical Positivism*, The Free Press, Nueva York, 1959, por donde cito; hay trad. castellana en el Fondo de Cultura Económica).
- «The Methodological Character of Theoretical Concepts», en Feigl y Scriven, *Minnesota Studies in the Philosophy of Science*, vol. I, University of Minnesota Press, Minneapolis, 1956.
- McCarthy, T.: «The Operation Called *Verstehen*: Towards a Redefinition of the Problem», en Schaffner y Cohen. *Boston Studies in the Philosophy of Science*, vol. XX, Reidel, Dordrecht, 1974.
- Natsoulas, T.: «Concerning Introspective 'Knowledge'». *Psychological Bulletin*, 1970.
- Nelson, R. J.: «Behaviorism is False», *The Journal of Philosophy*, 1969.
- Nisbett, R. W. y Wilson, T. D.: «Telling More Than We Can Know: Verbal Reports on Mental Processes», *Psychological Review*, 1977.
- Quine, W. O.: «On Mental Entities», *Proceedings of the American Academy of Arts and Sciences*, Boston, 1951, recogido en O'Connor, *Modern Materialism: Readings on Mind-Body Identity*, Harcourt, Brace and World, N. York, 1969, por donde cito.
- «Epistemology Naturalized», en *Ontological Relativity and Other Essays*. Columbia University Press, N. York, 1969 (hay trad. castellana en Tecnos).
- «Reflexiones filosóficas sobre el aprendizaje del lenguaje», *Teorema*, 1972.
- *The Roots of Reference*, Open Court, La Salle, 1974 (hay trad. castellana en la Revista de Occidente).
- Rivière, A.: «El análisis experimental de la conducta y el conductismo radical como filosofía», *Investigación y Ciencia*, abril, 1977.
- Rorty, R.: «Incorrigibility as the Mark of the Mental», *The Journal of Philosophy*, 1970.
- Ryle, G.: *The Concept of Mind*, Hutchinson, Londres, 1949. (Penguin, 1963, por donde cito; hay trad. castellana en Paidós).
- Sellars, W. F.: «Empiricism and the Philosophy of Mind», en Feigl y Scriven, *Minnesota Studies in the Philosophy of Science*, vol. I. University of Minnesota Press, Minneapolis, 1956 y reproducido en la obra del autor, *Science, Perception and Reality*. Routledge, Londres, 1963, por donde cito (hay trad. castellana en Tecnos).
- Skinner, B. F.: *About Behaviorism*, Knopf, N. York, 1974 (hay trad. castellana en Fontanella).
- Wittgenstein, L.: *Philosophische Untersuchungen*, Blackwell, Oxford, 1953.
- *Blue and Brown Books*, Blackwell, Oxford, 1958 (hay trad. castellana en Tecnos).
- «Wittgenstein's Notes for Lectures on 'Private Experience' and 'Sense Data'», *Philosophical Review*, 1968.

BIBLIOGRAFIA

(Sobre «Salud mental y sociedad»)

Por un error de transcripción, ajeno a este Boletín, en la Bibliografía del ensayo «Salud mental y sociedad», del profesor Florencio Jiménez Burillo, publicado en el número anterior, no se reseñaron las obras a las que precisamente se hacía referencia en el texto.