



6
24 03081

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

INSTITUTO DE INVESTIGACIONES BIOMEDICAS

**PROBLEMAS TEORICOS, CONCEPTUALES Y
METODOLOGICOS RELACIONADOS CON EL ESTUDIO DEL
COMPORTAMIENTO PROPOSITIVO**

TESIS
Para optar por el Grado Académico de Doctor en
Investigación Biomédica Básica

presenta

Guadalupe
NYDIA LARA ZAVALA

México, D.F. **TESIS CON**
FALLA DE ORICEN

1997



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

PAGINACION VARIA

COMPLETA LA INFORMACION

**A mis hijos
Aidin y Roberto**

AGRADECIMIENTOS

Agradezco la ayuda que me prestaron en la elaboración de este trabajo al Dr. Francisco Cervantes Pérez por haber dedicado tiempo y esfuerzo en la dirección, corrección y concreción de esta empresa. A la Dra. Carmen Gómez por su inapreciable amistad y confianza. Al Dr. Claudio Firmani Clementi por proporcionarme los espacios de desarrollo y crecimiento requeridos para la realización de ésta tesis. Al Dr. Felipe Lara Rosano tanto por su inmensa generosidad, como por haberme brindado la oportunidad de compartir mi trabajo a nivel internacional. Al Dr. Lino Díaz de León y al Lic. Jorge Pérez de la Mora por todas sus atenciones. Al Dr. Edmundo Lamoyi y a la Srita. Blanca Huerta por todo el apoyo que tan desinteresadamente me brindaron para lograr concretar la etapa final de mi trabajo. A los participantes del *Seminario-taller continuo sobre problemas teóricos, conceptuales y metodológicos de modelos computacionales inspirados en sistemas biológicos*, en especial a tres de sus más entusiastas colaboradores: Luis Mendoza Sierra, Francisco Caviedes Contreras y Gabriel García Acosta, por lo mucho que significaron sus intervenciones, críticas y comentarios para mi trabajo. A Diana Castillo García por todos sus esfuerzos para solucionar cualquier traba relacionado con mi trabajo. Por último, quiero agradecerle de manera muy especial a mi compañero y maestro Alejandro Tomasini Bassols el insustituible apoyo moral, emocional e intelectual que siempre me brindó a lo largo de todo el proceso.

INDICE

Introducción general	6 páginas
Forms of explanation	31 páginas
Methodological considerations in Cognitive Sciences	14 páginas
Methodological considerations in System and Cognitive Sciences	5 páginas
Conductismo y psicología cognitiva	8 páginas
Filosofía y ciencias de la conducta	14 páginas
Metafísica, procesos mentales y observaciones empíricas	20 páginas
Newton: percepción y explicación científica	16 páginas
Perception and scientific explanation	5 páginas
Percepción, metafísica y explicación científica	7 páginas
Emergence: a tension between definitional and empirical properties	6 páginas
Problemas conceptuales y metodológicos en torno a la noción de 'sistema'	3 páginas
System Science and the study of complex behavior	6 páginas
Conclusiones generales	5 páginas

INTRODUCCIÓN GENERAL

Las ciencias biológicas han generado sólo una gran teoría "omniabarcadora": la teoría de la evolución. Ahora bien, si dicha teoría tuviese las ya bien conocidas características de, digamos, una teoría física (formalización, axiomatización, predicción, etc.), entonces debería ser igualmente rica y absorber dentro de su marco teórico la explicación de prácticamente cualquier estudio de los sistemas biológicos. Empero, no es éste caso. La teoría de la evolución sólo ofrece un panorama general en torno al origen y secuencia histórica de la formación y evolución de los sistemas vivos. No obstante, la sistematización de los resultados hasta ahora alcanzados requiere de una teoría de la evolución más acabada o de una teoría alternativa.

Es relativamente claro, por otra parte, que el examen minucioso de una teoría científica de las magnitudes de la teoría de la evolución comporta el escrutinio de dificultades que rebasan el marco del trabajo de laboratorio y, en verdad, de la investigación estrictamente biológica. Son muchos los problemas concernientes al status, estructura, funcionamiento, etc., que plantea una teoría como la de la evolución. Es, por ejemplo, perfectamente posible (si bien altamente debatible) el que preguntas básicas (explícitas o implícitas) de la teoría de la evolución, como pueden ser preguntas acerca del origen de la vida (al igual que preguntas de cosmología abstracta o de lingüística) sean interrogantes desorientadores, o pseudo-preguntas. Por ello, podría por ejemplo sugerirse que de lo único de lo que puede ocuparse la teoría de la evolución es, a final de cuentas, no de "problemas" como los del origen de la vida, sino más bien y simplemente de diversos aspectos de su "desarrollo".

Puede sostenerse que, en principio, la teoría supone la existencia de una cadena causal que va de lo simple a lo complejo y que entre un sistema y otro no hay cortes, huecos, discontinuidades, etc., a pesar de que, por otra parte, se aceptan cambios bruscos generados por mutación. Pero, como es obvio, el concepto de mutación es sumamente problemático y es de difícil integración a una teoría que incorpora conceptos de lógica muy distinta. Esta clase de problemas no reciben una respuesta o solución obvia por parte de los teóricos de la evolución. El esclarecimiento de estos problemas requiere de una fuerte dosis de filosofía de la ciencia y del lenguaje.

Que la teoría de la evolución está expuesta a objeciones serias es algo que fácilmente puede apreciarse. Por ejemplo, puede defenderse la tesis de que si la teoría explicara el desarrollo de los sistemas biológicos y diera los lineamientos para seguir su secuencia histórica, su trayectoria, etc., en principio no se tendrían problemas ni para clasificar las especies ni para explicar en forma global los distintos aspectos que se pueden estudiar en los distintos seres vivos (e.g., molecular, conglomerado molecular o agregados moleculares, células, tejidos, órganos, sistemas, poblaciones, etc.). La teoría, sin embargo, no contiene u ofrece una metodología general para estudiar los distintos aspectos de los sistemas biológicos y sus componentes. La prueba de ello es que diversas áreas de la misma disciplina tienen que recurrir a diferentes técnicas de investigación y utilizar diferentes conceptos y herramientas metodológicas para conseguir sus resultados. Así, a pesar de que las diferentes ramas de la biología requieren de un instrumental científico

muy preciso para generar sus mediciones, las interpretaciones que se generan de esas mediciones rara vez pueden ser integrados en un todo coherente. Esto se debe, en parte, a que la mayoría de los resultados obtenidos a través de esas mediciones tienen tan sólo una interpretación *ex post facto*. Esto significa, que los experimentos que se realizan se diseñan para causar una perturbación en el comportamiento de un sistema y que es en función de lo que le acontece al sistema así estudiado que se genera una posible interpretación del suceso. Puede, pues, inferirse que en biología, conforme aumenta el número de datos acumulados gracias al estudio de diversos sistemas orgánicos, aumenta el número de explicaciones *ex post facto*. Los nuevos datos difícilmente sirven para confirmar o rechazar los postulados de una teoría *a priori* en torno a la vida, su génesis y su desarrollo. Por esta razón, las explicaciones biológicas constantemente aparecen como un conjunto interminable de piezas sueltas de un rompecabezas siempre incompleto. Esto tiene como efecto el que el biólogo tienda a pensar que lo único que le hace falta a la biología son nuevos experimentos, para generar nuevos datos. Pero este enfoque, de hecho adoptado por muchos biólogos, es, por lo menos, cuestionable y poco "rentable".

El problema general se puede resumir afirmando que, mientras el biólogo siga mecánicamente generando datos que sólo puedan ser explicados *ex post facto*, el rompecabezas que es actualmente la biología seguirá inevitablemente apareciendo como una ciencia esencialmente incompleta e inconexa. Sin embargo, es factible defender el punto de vista de acuerdo con el cual no son nuevos datos los que se requieren para comprender el fenómeno de la vida y su comportamiento, sino que lo que se necesita es algo más fundamental, a saber, una teoría que además de ser capaz de describir los hechos observados, ofrezca junto con ella un método y un formalismo lo suficientemente rico, de manera que la investigación quede orientada de manera definida y que la totalidad de los resultados queden articulados en un único sistema coherente. Para ello, empero, se necesita conjugar el conocimiento y el manejo del instrumental conceptual y formal indispensable para aprovechar de manera efectivamente fructífera los resultados de lo que sería una genuina biología teórica.

Es menester enfatizar que, dada la variedad y complejidad de los sistemas biológicos y la carencia de un método y un formalismo general propio de la biología, el estudio (en toda su variedad) de los seres vivos hace que se recurra de manera un tanto caótica a toda una diversidad de métodos y de herramientas formales. Así, por ejemplo, algunos estudios biológicos utilizan los conocimientos de la física y la química junto con sus herramientas formales para modelar y explicar ciertos procesos y mecanismos observados en los seres vivos. Otros, sin embargo, prefieren utilizar el formalismo de las ciencias de la computación para generar sus modelos. Y aun otros, utilizan de manera intuitiva rasgos de la teoría de conjuntos para generar clasificaciones taxonómicas.

Esto ha ocasionado que proliferen distintos modelos biológicos, apoyados todos ellos en distintas áreas de conocimiento, y que utilicen distintos sistemas formales para elaborarlos (matemáticas, lógica, lingüística, teoría de la computación, teoría de la información, etc.). Esto mismo ha llevado a la biología a dividirse en infinidad de áreas de investigación que sólo de manera parcial explican las características y los mecanismos a través de los cuales operan los sistemas biológicos.

Parte de las dificultades que confronta actualmente la biología es que este proceder desordenado impide sistemáticamente la integración de los distintos modelos, así como la gran variedad de resultados obtenidos en lo que de hecho es una enorme gama de disciplinas que dan lugar al *corpus* del conocimiento biológico. Eliminar el desorden en el que se encuentra la biología exige, pues, de una revisión general no sólo de modelos biológicos concretos, sino de un examen minucioso de su lenguaje, conceptos, métodos, tipos de formalismos, etc., pues sólo un estudio así podría permitir explorar seriamente la posibilidad de la unidad teórica y metodológica de la biología actual.

Ahora bien, la constitución de una biología teórica útil, no sólo requiere de la participación de gente entrenada en biología, sino que hace falta la participación y el aprendizaje del conocimiento que se genera en otras áreas (*i.e.*, filosofía, psicología, sociología, lingüística, matemáticas, ingeniería, ciencias de la computación, etc.). Este fue el proyecto que inició la gente de Biomédicas al aceptar en sus programas de maestría y doctorado en Investigación Biomédica Básica individuos con una formación distinta a la biológica. El problema aquí es que al hacer intervenir a gente formada en otras disciplinas en el proceso de aprendizaje y formación de un equipo capaz de confrontar los problemas a los que tiene que enfrentarse un biólogo teórico, la información y las cuestiones a tratar pueden aparecer como la exploración de problemas que nada tienen que ver con el objetivo principal. Expliquemos brevemente por qué analizando un caso particular: el mío.

Yo entré a estudiar la maestría y el doctorado en Investigación Biomédica Básica con especialización en Biomatemáticas viniendo del área de filosofía. El proyecto de investigación que propició mi entrada al doctorado consistía en explorar, desde el punto de vista teórico, conceptual y metodológico, si había manera de limar las asperezas que se generan entre diversas disciplinas interesadas en entender y modelar el denominado 'comportamiento propositivo' que caracteriza no sólo a los seres vivos, sino a todos aquellos sistemas (naturales y artificiales, *e.g.*, computadoras, robots, sistemas de control, redes neuronales artificiales, etc.) que realizan una función, meta o tarea concretas. Esto me llevó a la necesidad de generar una búsqueda más cercana hacia lo teórico que a lo experimental, aunque han sido los datos extraídos de la biología experimental y de las aplicaciones concretas de las ciencias de la computación lo que me ha servido para teorizar en torno a los distintos problemas que dividen, frenan y antagonizan, interna y externamente, a las disciplinas que estudian y modelan el comportamiento propositivo.

A lo que me enfrenté en el desarrollo de mi investigación fue al hecho de que la gran mayoría de los problemas que afectan no sólo la posibilidad de colaboración entre disciplinas que estudian lo mismo, sino la de reunir los datos y resultados obtenidos en una misma área de conocimiento, así como el sano desarrollo de diversos campos de investigación se debe a que:

- a) datos y resultados obtenidos en una investigación concreta se interpretan y explican apoyándose en alguna teoría de corte metafísico que no necesariamente es la misma que gente investigando lo mismo o algo parecido acepta *a priori*.
- b) la gente esta tratando de estudiar en cerebros y máquinas los mecanismos de procesos metafísicos, como son, por ejemplo, los denominados 'procesos mentales' (*e.g.*, percepción, inteligencia, pensamiento, etc.).

- c) se está utilizando el modelo explicativo de la denominada 'teoría causal de la percepción' para justificar la búsqueda de mecanismos metafísicos.
- d) doctrinas metafísicas de corte materialista emplean el lenguaje de las ciencias físicas para dar cuenta de elementos y procesos que la física ni pretende ni puede explicar.
- e) se están utilizando formas explicativas problemáticas, a saber, la teleológica y la mecánica, para entender el comportamiento de aquellos sistemas que llevan a cabo una función, meta o tarea concretas, es decir, lo que aquí estoy llamando 'sistemas propositivos'.
- f) la metodología que se utiliza para estudiar los mecanismos del denominado 'comportamiento propositivo' está siendo empañado por el uso de estas formas explicativas problemáticas.

El punto (a) lo desarrollo fundamentalmente en los trabajos que llevan los títulos *Methodological Considerations in Cognitive Science* y en *Methodological Considerations in System and Cognitive Sciences*. El segundo trabajo surgió del primero, pues buscando el uso de la noción de 'sistema' en las diversas áreas que estudian lo que se considera conducta "inteligente", me di cuenta de que lo que en el primer trabajo se estaba tratando como si fueran dos corrientes más o menos afines de las denominadas 'ciencias cognitivas', en realidad eran dos grupos interdisciplinarios que, aunque en apariencia investigaban "lo mismo", no sólo no partían de los mismos principios, sino que eran grupos claramente antagónicos: el de ciencias de sistemas y el de ciencias cognitivas.

El punto (b) está presente en todos los trabajos, pero dada la variedad de disciplinas involucradas se optó por revisar áreas concretas en los ensayos que llevan los títulos *Conductismo y psicología cognitiva*, *Mente-cuerpo y ciencias cognitivas* y *Metafísica, procesos mentales y observaciones empíricas*. El primero, como su mismo título lo indica, contrasta algunas de las propuestas del conductismo con las de la psicología cognitiva. El segundo es una revisión en tomo a la relación que históricamente se ha establecido entre la lógica, la filosofía de la mente y la filosofía del lenguaje. El tercero es una revisión crítica de trabajos concretos realizados por diversos reconocidos neurofisiólogos que pretenden que es posible observar procesos mentales a través de la actividad neuronal.

El punto (c) lo desarrollo de manera puntual en los trabajos que llevan los títulos *Newton: percepción y explicación científica*, *Perception and scientific explanation* y *Percepción, metafísica y explicación científica*. En el primero expongo la crítica que hace Newton a la teoría causal de la percepción para ofrecer un enfoque alternativo a dicha teoría. En el segundo trato de esclarecer el concepto y uso del término 'percepción' con el propósito de rastrear de dónde viene la idea de que la percepción es un proceso mental. En el tercero trato de contrastar lo que efectivamente puede estudiar un científico y lo que pretende estudiar el metafísico cuando analizan la percepción de un animal y sus problemas.

El punto (d) también se toca en varios trabajos, pero en *Emergence: a tension between definitional and empirical properties* es donde desarrollo cómo se establece la

liga entre el lenguaje científico y el metafísico, así como la distinción entre lo que el científico quiere y puede explicar y lo que el metafísico quiere y no puede explicar.

El punto (e) lo desarrollo ampliamente en el trabajo que lleva el título de *Forms of explanations*. Este trabajo surgió de la necesidad de definir con claridad el contraste, por no decir antagonismo, que se genera entre explicaciones mecánicas y teleológicas que, dicho sea de paso, están presente en todos los estudios relacionados con el comportamiento propositivo. Este tema originalmente se planteó como un pequeño apartado para completar el trabajo titulado *System science and the study of complex behavior*. Sin embargo, al revisar históricamente la génesis de estas dos formas explicativas pude darme cuenta de que es debido a la tensión que se genera entre ellas lo que ocasiona que se introduzcan posiciones metafísicas para solucionar las inconsistencias que surgen cuando las explicaciones mecánicas tratan de absorber los contenidos mentales que caracteriza a las teleológicas. Evidentemente, por la importancia del tema, decidí hacerlo no sólo un trabajo independiente, sino incluirlo como el trabajo que introduce la problemática que se trata en los demás. Por esta razón, *Forms of explanations*, aunque es el último de mis escritos, aparece al principio de esta tesis.

Finalmente, el punto (f) lo desarrollo en dos trabajos. El primero es el que se titula *Problemas conceptuales y metodológicos en torno a la noción de 'sistema'* y el segundo es el de *System Science and the study of complex behavior*. En el primero expongo la posición metodológica de Ludwig von Bertalanffy frente a los conflictos suscitados entre vitalistas y mecanicistas, así como la manera como Norbert Wiener introduce el concepto de procesamiento de información y con él una serie de confusiones para dar cuenta del comportamiento propositivo. En el segundo trabajo trato de explicar la metodología que se utiliza para estudiar el comportamiento propositivo, así como los malos entendidos que se producen cuando se trata de buscar dentro del sistema la causa de su propósito.

Evidentemente, para poder detectar y elaborar el contenido de cada uno de estos problemas fue necesario invertir mucho más tiempo de lo pensado. Esto se debió, fundamentalmente, al hecho de que un trabajo de esclarecimiento teórico, conceptual y metodológico en torno a las interpretaciones científicas carece de antecedentes en nuestro tiempo. Los filósofos (e.g., Paul Churchland, Patricia Smith Churchland, Daniel Dennett, Hilary Putnam, John Sarle, Fred Dreske, Karl Popper, Jerry Fodor, Mario Bunge, Noam Chomsky) que se han vinculado al estudio de los problemas metafísicos, que sin duda están presentes en prácticamente todas y cada una de las disciplinas que estudian el comportamiento propositivo, los han trabajado defendiendo la perspectiva de una doctrina metafísica particular. Esto se debe a dos factores importantes que no sólo son dignos de ser mencionados, sino que es necesario combatir. El primero es que el filósofo tiende a pensar que las interpretaciones científicas se generan única y exclusivamente atendiendo a los datos y resultados que se desprenden de su investigación. Desde esta perspectiva, es claro que se piense que si la comunidad científica acepta esos datos y resultados no es competencia del filósofo cuestionarlos. El segundo es que, cuando el filósofo encuentra en la literatura científica la defensa de posiciones metafísicas semejantes a la suya, cree que lo que el científico le está proporcionando son los fundamentos empíricos de determinada doctrina filosófica.

Obviamente, lo que el filósofo no considera en el primer caso es que aunque los datos y resultados que proporciona el científico pueden estar completamente fuera de duda, eso no significa que no sea posible cuestionar la interpretación que el científico ofrece de esos datos y resultados. En lo que refiere al segundo caso, lo que no se está tomando en cuenta es que cuando una pseudo-pregunta se enreda en el análisis científico, normalmente lo que sucede es que el científico se ve materialmente forzado a adoptar, consciente o inconscientemente, una determinada posición metafísica sobre la naturaleza de la materia, de la vida, de la mente o del lenguaje, para lograr darle coherencia a lo que quiere demostrar con sus datos y resultados.

Ante esta consideración podemos decir que, si el filósofo sigue pensando que su labor consiste en justificar una necesidad metafísica en la ciencia, apoyándola con los datos y resultados que ésta le proporciona, evidentemente sería más sano para la ciencia que el filósofo se mantuviese al margen de su quehacer científico, pues con su actitud lo único que logra es complicar más la comprensión del obstáculo que le dio a la ciencia cabida a la especulación filosófica. Sin embargo, existe otra alternativa.

Los doce trabajos que estoy presentando como tesis doctoral lo que tratan de hacer es justo revisar cómo y por qué posiciones metafísicas se entremezclan en la interpretación de los datos y resultados científicos. En todos ellos de alguna manera trato de explicar el origen de las pseudo-preguntas que gestan la dañina tendencia de buscar en la metafísica algún tipo de respuesta a preguntas que distan mucho de ser genuinas. Asimismo, en muchos de ellos trato de ofrecer posiciones alternativas para evitar que el quehacer científico se distraiga, por no decir se empobrezca, buscando cosas que el científico ni va a poder encontrar, ni le compete buscar. Si esta actitud efectivamente puede darle claridad al científico y enseñarlo a distinguir entre los genuinos problemas de la ciencia y las cuestiones que sólo lo pueden conducir hacia los pantanosos terrenos de la metafísica, entonces creo que la labor filosófica no sólo queda plenamente justificada, sino que, independientemente de cuánto logre aportar yo con mis trabajos, la filosofía puede ser vista como un arma teórica y conceptual indispensable para el sano desarrollo no sólo de la biología en particular, sino de la ciencia en general.

FORMS OF EXPLANATION

Nydia Lara
Laboratorio de Neurocomputación
Centro de Instrumentos
UNAM

INTRODUCTION

Most people think that science must provide us with causal explanations of natural phenomena, but most people does not realize that: first, the term 'cause' has more than one meaning; second, their different meanings give rise to special forms of explanation; and, third, there are some forms of explanation which can be properly considered as scientific, while others are the result of metaphysical speculations, that is, foreign to science.

Our aim in this work is to review some of the most common forms of explanation used by different disciplines, in order to show not only its origins, but the realm of reality, or unreality, that might be explained by using them. Seven forms of explanation are described and discussed: four provided by Aristotle, two by the metaphysical thinking, and the last one by Newton. For every forms of explanation we review how they have been historically conformed and why metaphysical forms of explanation have been mistaken or mixed with those considered as scientific. To elucidate this last idea, we pay special attention to the review of the two most common and problematic metaphysical forms of explanation: the teleological and the mechanical, which are important because they have been used not only by biologists, but also by most professionals belonging to those disciplines concerned with the understanding and explanation of the so called '*purposive*' or '*teleological*' behavior. Our objective is to show that teleological explanations have to introduce mental terms in order to explain the causal source of that kind of behavior, whereas mechanical explanations have introduced the mind-body distinction, in addition to the absurd attempt of explaining in physical terms the causal source of mental processes (e.g., purposes, desires, intentions, perceptions, etc.) This last attempt indeed has been affecting the healthy development of biological sciences, in particular, that of the neurosciences.

In addition, we analyze how the attempt to give the physiological basis for the so called 'mental processes' is the conceptual result of the most bizarre metaphysical theories ever conceived, and that those constructions are the result of a conceptual mistake that has its origin in the Christian notion of efficient cause.

The notion of efficient cause was originally coined by Aristotle to explain the modification, alteration or change of what he calls a 'natural process'. The conceptual scheme of the Christian notion of efficient cause was taken from Aristotle, but the Christian used this form of explanation aim to explain, besides the modification, alteration or changes of natural processes, the causal source of natural processes themselves. To understand how the search for the explanation of natural processes is affecting the interpretation of the data and results

not only in biology, but also in psychology and even in physics, we start our review by trying to clarify, in the next sections, the different meanings of the term 'cause'.

1. Aristotle

According to Aristotle all kinds of explanations of natural phenomena are causal explanations. He distinguished between four explanatory causal principles: the material, the formal, the efficient, and the final. Material causes deal with the explanation of substrates; formal causes with substrates' shape; efficient causes with what initiates changes in terms of sources; and, final causes with the end, or purpose, for which a particular change or process takes place. That is, material and formal causes are used to explain processes and changes intrinsic to the objects, efficient causes are concerned with processes and changes extrinsic to the objects, and final causes pretend to explain the end or result of a change or process.

Every form of explanation has its own realm of knowledge, and Aristotle criticized his predecessors for emphasizing some of them and neglecting others. The Milesians, for example, were concerned too much with material causes; they believed that the whole world and its different manifestations could be explained defining only the nature of the basic matter of which all things are composed. The concept of an efficient cause emerged with Empedocles, who try to account for all changes and transformations of the world only in terms of acting forces. Plato often spoke as if the explanation of all things could be achieved simply by discovering their forms (formal causes). Nevertheless, for Aristotle, all these procedures were highly limited, in the sense that each causal principle by itself could only explain part of the story (Aristotle, *Metaphysics*, Book I). To illustrate how each causal principle relates to others, Aristotle provide us with a simple example:

...the moving cause of a house is the art of the builder (the efficient cause), the final cause is the function it fulfills, the matter (the material cause) is earth and stones, and the form (the formal cause) is the definition (Aristotle, *Metaphysics*, Book III, Ch. 1-2, 996b).

It is clear that the knowledge of earth and stones, by itself, could never give us the knowledge of the formal, final or efficient causes of a house. In the same way, our knowledge of the forces needed to build a house, will never allow us to know anything about its material, formal or final cause, neither the knowledge of the house plane would tell nothing about its material, efficient and final causes. However, the material and formal causes are hard to separate in the Aristotelian way of thinking because, even when he recognized in his example that earth and stones are to be understood as the component parts of the house, earth and stones have their own formal causes. This was possible because matter, for Aristotle, is the substratum capable to acquire a form, and a form is, accordingly, what gives the substratum its being. That is, for Aristotle there are no empirical knowledge of the substratum without a form, and no empirical knowledge of forms without a substratum. In that way, both, matter and form are not ontological entities that have an existence of their own, rather they are explanatory principles.

Matter might acquire any form, those, for Aristotle matter is potency; whereas form is the definition of matter, and in Aristotle's theory it is defined as an actualities (instances) of

matter, that is, it is what gives its shape to particular objects. Substances are not entities, they are names defining a class of entities that share the same form. Changes and processes consist of a substrate (matter) which has acquired a form that it did not previously possess (G.B., Kerferd, Encyclopedia of Philosophy, Vol. I, Aristotle, p. 156).

Efficient and final causes, in their turn, aim to explain opposite things. The first explain what provokes or initiate a change or a process, whereas the second, explains (*a posteriori*) the end or result of that change or process (Aristotle, Physics, B.II, ch. 3, 195a (30)). Changes or processes produced by efficient causes always require physical agents. On the contrary, purposes in the sense of ends, are conceptual devices that only pretend to explain the final result of something that has been done. Final causes, therefore, are not linked with agents (neither physical nor mental) and, at least for Aristotle, it was clear that final causes should never be mistaken by efficient causes. For him, a house builder is indeed the appropriate explanation of the efficient cause of a house, but the house builder is not relevant for explaining the purpose of the house (its end or final cause). The purpose of the house is clearly independent of the agent that makes it possible to build the house, as a matter of fact, our knowledge of final causes is not related, or depending, on our knowledge of efficient causes.

Taken all together, we may say that with the four causes doctrine, Aristotle tried to explain not four, but only three different ways to understand changes and processes: the first tries to explain changes and processes in terms of how matter acquires a specific form; the second tries to explain how changes and processes are produced by an agent; and the third seeks to explain the final state or the purpose that any changes or processes achieve. It seems that Aristotle considered the first as our own way to grasp the knowledge of a natural processes; the second as our understanding of a change or process that, in order to start, it is assumed to require the intervention of a force (an agent); and, the third as our way to intellectually grasp the purpose associated to the change or process that took place, that is, its end or final result. If this is correct, it is possible to assert that Aristotle faced two kinds of changes and processes to be explained: the spontaneous or natural, and those produced by agents. Agents are understood as forces that change, disrupt or modify spontaneous process, or as forces that had to intervene to give matter a form that it will not possess or acquire spontaneously.

Aristotle sustained that in nature there are many spontaneous or natural changes and processes that can be observed. Growth being one of them. As a biologist, Aristotle tries to account for them in terms of his theory of potency and actuality. According to his point of view, if some conditions are given, and no external agents intervene, then an oak seed, for example, "spontaneously" is going to grow as an oak tree. The process of growing can be explained in terms of a material cause (the seed) acquiring a form (of a tree). In his terms, the seed is an oak tree in potency, and the oak tree is simply the actuality of the seed. He calls the spontaneous or natural transition from potency to actuality an '*entelechy*' or '*natural teleology*'.

Now, clearly for Aristotle, in order to explain entelechies or natural (teleological) processes no efficient causes or agents are required. Our previous knowledge of the material (the

seed), the form (the oak), and the notions of potency and actuality is sufficient to give a complete explanation of the whole process. That is, the knowledge of material, formal and final causes are then enough to let us explain and predict what is going to happen with an oak seed if we previously know the final result of the process and some of the conditions that have to be given in order to allow the transition of a thing from potency to actuality. However, something different happens when changes and processes are produced by agents. A piece of marble, for example, even when in potency it is able to become a statue, it will not take that form unless an efficient cause intervenes (a sculptor). The sculptor (that in this case plays the role of an efficient cause) might be able to "force" de marble (the material cause) to acquire a specific form. Nevertheless, without the sculptor it is obvious that the marble by itself will never become a statue. It is the sculptor who has to give the marble the shape (its formal cause), and the result of the effort and skills of the sculptor will give the statue its final cause (its beauty).

Notice that, for Aristotle, all changes and processes (spontaneous or produced) are explainable in strictly physical terms. These take us to our last remark. For Aristotle all forms of explanation need, as its point of departure, a material cause (*i.e.*, a substratum). However, it is not the substratum what is going to be explained with the other causes, but those changes and processes that the substratum might undergo. In Aristotle's doctrine, it is the substratum what acquires a form and the one that is susceptible to be transformed. We are able to explain the transformations of the substratum in terms of natural (*entelechies*) or produced causes (efficient causes). Nevertheless, the substratum for Aristotle does not work as an ontological principle. It works as an explanatory principle, that is, as a general proposition with the power to link our explanations of changes, processes and transformations with the world we perceive daily.

Nevertheless, when the Aristotelian thought became allied with Christian theology his theory of the four causes was altered and reinterpreted in the light of theological ideas. Many of those ideas had, for centuries, as their second nature Augustinian and Platonic interpretations, and the philosophers of the late Middle Ages fashioned with them their own "Aristotelian" version of the four causes. Unfortunately, it is their Aristotelian version the one that has survived up to our time, and the main problems behind the modern discussion between teleological and mechanical explanations are derived from the way theologians reconstructed Aristotle's theory of the four causes. In the next section we briefly review what they did, in order to understand the difficulties that arise with this "new" treatment of the four causes.

2. Aristotelism

Christian theologians were not particularly interested in explaining those changes and processes we observe daily. They were more interested in cosmological issues related to the genesis and the source of the universe. They, in fact, wanted to account for the creation of everything, including matter, shape (form), and the purpose of the universe only in term of one cause: God. They only needed the explanatory power of efficient causes to put God as the agent that created the universe. They also linked the notion of efficient cause with the

idea that "nothing could be what it is without the intervention of an agent". This statement is the one that conforms till today the basic idea of what is known as the 'causal principle'. Obviously, the use of the term 'cause' here suffers a great alteration. 'Cause' in Aristotle's sense means a conceptual form for explaining what exist. He did not conceived causes as elements of existence themselves. Even when efficient causes required physical agents (existing agents, e.g. natural or artificial factors, instruments, human beings, etc.) terms like 'force' and 'skills' were only introduced as one way for explaining the occurrence of an event. Causes in all cases play for him a conceptual role, rather than an ontological one, and, in that sense, were only explanatory principles with the conceptual power to account for those changes and processes that can be observed around us. However, Christian theologians postulated God as the first existing element of a causal chain. Hence, they wanted God to be an ontological cause, and not only an explanatory principle. Let us try to understand the consequences of this conception.

Following some of Plato's ideas, Christian theologians were concerned with constructing a cosmological argument to ensure God's existence. The argument proceeds from the fact that the universe exist. Following this initial premise they simply applied the idea that everything that exists must have a cause. The universe exist and, hence, it must have a cause, which happens to be God, therefore, God must exist (Hospers, p.429, 1967). This argument clearly allowed them to interpret the existence of the sensible world as the effect of God's actions, which they wanted to be an ontological cause (For an interesting review of the conceptual problems that arise with this idea, see Tomasini, *Filosofia de la Religión*, 1996). From the premise that the existence of the universe is an effect, they could properly infer the existence of the cause (God). The "inferred" cause, in this case, evidently does not work as an explanatory principle, but rather as the logical and ontological condition for the existence of what they *a priori* considered to be an effect (the universe). This means that the inferred cause by definition becomes to be the most important existing element to explain the existence of the effect. Clearly, God, as an ontological cause, enables them to advance a step beyond the sensible world, and let them explain everything that exists as an effect. In addition, they went a step further, and explained the existence of everything in the universe by having God as their cause, that is, through a non physical cause.

Under this conception, formal causes were interpreted in accordance with the Platonic Ideas; that is, as archetypes that, being the objects of the divine intellect and the patterns of creations, represent God's thoughts and the measures of "God's truth". Aquinas introduced the term 'truth' as "the adequation of thing (*res*) and understanding". However, the application of that definition varies. In the case of human beings, truth were properly attributed to our understanding of things, whereas in the case of things, they were true as long as they correspond to God's thoughts (Warner Wick, *Encyclopedia of Philosophy*, Vol.1, *Aristotelianism*, p. 150). Natural processes embody God's plan; hence, Aristotle's distinction between natural and produced events simply vanishes. All events are, according to this scheme, produced by an efficient cause. As we will see below, this idea has to be considered responsible for all kinds of metaphysical speculations about the causal source of what had been previously considered by Aristotle to be as natural processes.

For Christian theologians final causes, instead of being understood as our way to grasp the final result of a change or of a thing done, were conceived as the fulfillment of God's plan. Under this interpretation, a great step toward the unknown was given, because the whole conception accepts, somehow, that non physical causes (e.g., plans, goals, desires, etc.) can work as efficient causes in its ontological sense. Since then, final causes became understood as something carry out with a purpose, where the term 'purpose' does not mean the conceptual interpretation of the result of a change, but rather the ontological cause of a change executed according to an agent's plan, intention, desire, etc. This idea was clearly far from Aristotle's understanding of final cause. We shall call this mixed version of final and efficient causes as '*teleological explanations*'. Obviously, this sort of explanations not only reduce the four causes offered by Aristotle into an altered form of efficient causes, but we have to be aware that this distortion represents both, a different conceptual scheme, and a different way of understanding the term 'cause'. Before we can discern the harm this new form of explanation has made within all those disciplines that have to deal with purposive behavior, in the next section we briefly review the transition that the concept of God suffers after the Middle Ages.

3. Theism and Deism

The Middle Ages doctrine of creation is primarily concerned with the ontological dependence of the world upon God. In that sense, God was not conceived only as a builder that once he finished his work can go away. God, according with this doctrine, has to sustain the world, and as Saint Augustine puts it, if God were to cease supporting the world, even momentarily, it will altogether cease to be (Ronald W. Hepburn, *Encyclopedia of Philosophy*, Vol.2, Religious Doctrine of Creation). Now, the idea that God continues to exercise his influence over the course of the world at every stage of history is called '*theism*' (Hospers, p.427).

A theist conception of God is, at first sight, very comfortable. God is seen and understood as a supernatural, omniscient, omnipotent, and benevolent being that continuously is engaged with the care of his creatures. He hears and answers our prayers, he knows our needs and takes care of them. If we ask, or need, God to intervene in the course of any event, he can do it if he wishes. The problem with this conception of God is that science is futile. We would not be able to avoid or prevent maladies, even if we know some of the regularities of nature, because God can alter, reverse, adjust or modify them any time he wants to. A scientific knowledge (*i.e.*, the systematic observations of certain repetitions or regularities in nature), under the theist conception of God, would never be enough to predict the present and future course of nature. Anything might happen if God wishes to. Hence, the first feeling of comfort that we had with the theist conception of God, could become a feeling of discomfort when we want to predict or prevent future events. Deism can be seen as an expression of the discomfort that some of the theist formulations about the role of God embrace.

Deism begins when Christian theology accepts that God not only created the universe, but created us as intelligent beings with the capacity of knowing and working with the created

matter in order to be self sufficient. If that was the case, it was not clear why God has to be conceived as a being that all the time has to look after us. It was a lot more rational to think of God as a being who created the world as a splendid machine and then left it to run by itself. This idea is, as we will see below, the source of the mechanical thinking in its two versions: the philosophical and the scientific.

Deism flourished during the Renaissance and with some modifications little by little become the accepted creed of many philosophers and wise man of the Enlightenment. It endures till the late XIX century, and among its followers we find a variety of eminent persons like Montaigne, Descartes, Gassendi, Voltaire, Boyle, Locke, Newton, Clarke, Rousseau, Leibniz, Kant, Franklin, Jefferson and Washington. For these people, the term 'deism' embraces the idea that there is a God, who, without doubt, has to be considered as the first cause of the universe. In that regard deism and theism coincide. Although, deism differs from theism in the idea that the world God created was not only perfect, but also self sufficient, self sustained and self ruled (Ernest Campbell Mossner, *Encyclopedia of Philosophy*, Vol.2, Deism). In that respect, the most important legacy of deism is that, even when the doctrine accepts without hesitation that God was there behind the scene as its first cause, God was not the subject of our search for knowledge, but rather, it is his creation. More specific, the concern of deism is to know the rules that govern the movements and changes of what they conceived to be that huge machinery of the created universe. This step was important because, on the one hand, the deism thesis allows us to think that we live in a well ordered world susceptible to be known by us. On the other hand, religious considerations about the world could be separated, in principle, from our rational and empirical research of the world. This detachment took time, but it was deism that allows scientists and their adepts to be, or not to be, religious without altering the results of the research concerning the knowledge of the world we live in. Unfortunately, the legacy of the Christian thought about causation was strong enough that more often than not we are unable to distinguish the difference between genuine scientific knowledge and metaphysical speculations about the causes of phenomena occurring in our surroundings. This confusion is possible because, in the first place, we tend to associate the term 'cause' with a great variety of meanings, and, in the second place, the ambiguity of the term somehow tolerates the translation of a genuine scientific question into a metaphysical one. In the next section we try to understand how this translation is possible.

4. Causal explanations

All deist accepted that nothing could happen without a cause, but not all deist understood by the term 'cause' exactly the same. For some the term 'cause' always refers to the explanation of the source or production of an event, or of an specific behavior. Others, based in the idea that God created a well orderly world, introduced the idea of natural laws and employ the term 'cause' as a synonymous of the term 'law'. Laws, in contrast with the other and more common use of the term 'cause', never pretend to explain the source or production of anything. Laws have to be understood as general statements with no other purpose than to describe as precisely as possible, in theoretical or empirical grounds, the principles or rules

that are able to explain the observable occurrence of those events that usually display a regular pattern of behavior.

Obviously, the description of a regularity is not equivalent to the explanation of the source or production of an event. The first simply asserts how things behave spontaneously or in interaction. The second, by contrast, is never satisfied with the accurate description of an event. Its tendency compels people to obtain a fuller understanding of an observed regularity through the "knowledge" of the source that produces it. Nevertheless, the search for sources usually takes people to look for unknown metaphysical agents that in principle operate behind the phenomena (Rudolf Carnap, *An Introduction to the Philosophy of Science*, p. 12, 1966). We are, hence, in the presence of two different forms of explanation. The one that searches for laws, which gives rise to what we properly call a 'scientific explanation', the other that searches for the sources of those laws or regular pattern of behavior, which gives rise to metaphysical explanations.

Now, the dissimilarities of concerns derived from the divergent meanings of the term 'cause', somehow favors the interaction between scientific research and metaphysics. This is so, because it is always possible to attach a hidden metaphysical agent to explain what is supposed to be "behind" the regularity that scientists describe. However, metaphysical speculations about hidden agents, usually have the power to distract the attention of scientists and make them believe that the role of science consists in explain both, the regularity of the phenomena and its source. The problem is that metaphysical agents, by definition, are no accessible to the scientific method, and, hence, they add nothing to the scientific research, only deviates its concerns into the search for something that they will never find. In the following section we try to understand how metaphysical explanations operate and how they get involve with the scientific research.

a) Metaphysical explanations

Those who associate the term 'cause' with the idea of source or production, usually follow the Christian conception of efficient causes. This conception involves some peculiarities that are important to elucidate. One of them is that this notion of efficient cause makes us believe that everything that happens in the universe has to be considered as an effect, and, if that is the case, then, this conception necessarily leads us to believe that all things, facts, events or situations, that we observe around us must have a cause producing them. Other peculiarity is that when we assume *a priori*, that observable things, facts, events or situation are effects, obviously we have to consider, also *a priori*, that the cause that produces them must be behind of what we are able to observe. This means that, under this conception, causes are always hidden. Hence, causal explanations of this sort are necessarily associated with metaphysical speculations not only about the nature of those invisible or hidden causes, but also with the unavoidable question concerning its place in nature and the way they are able to operate behind the perceived world.

Now, the problem to define the nature, place and performance of those hidden causes has giving rise to all kinds of metaphysical doctrines. Some claim, for example, that the nature of

these causes are not physical but mental; others claim that they have to be understood as "special" qualities of matter. For some those "special" qualities are mental qualities intrinsic to matter, for others they are emergent properties of matter. Some claim that causes act like mechanical powers that "force" things from behind to behave in certain ways when they are present. For others causes are more like forms (e.g., programs, instructions, codes, etc.) that somehow dictate (or inform) matter from behind exactly what to do in each concrete situation. Some place them in the physical world as hidden or occult qualities, forces, mechanisms or forms. Others takes them out of the world we perceive, and assume that the physical world is not only outside our perception, but the hidden cause of the so called 'mental' or phenomenal world.

Obviously each metaphysical doctrine claims to have the correct answer to each one of these questions, but despite the variety, and sometimes contradicted or even opposite answers they offer to us, all agree in one point: that independently of how each doctrine decides to confront the problems related with the nature, place and ways of operation of those causes, they themselves do not form part of the perceived world. Hence, somehow, all metaphysical doctrines recognize they have not empirical grounds to decide which metaphysical answer is better than the other. Some might be more convincing than others, however, it is never possible to confirm or refute a metaphysical answer with the scientific data and results. The reason is obvious: metaphysical doctrines simply use scientific findings to try to explain something that is believed to be always and forever behind obtained data and results.

Newton and his followers were probably the first to realize that it was metaphysics who needed the scientific knowledge to convince people about the plausibility of its answers, but no the other way around. It most be pointed out that scientific knowledge, that is, the systematic study of the regularity of natural or experimental processes, can be obtained independently of the metaphysical concerns about the source of those regularities described by the scientists.

In fact, what distinguish science from metaphysics is that metaphysicians assume *a priori* that all observed phenomena have a cause that bring about, bring forth, create, or produce them. In this regard, they believe that an explanation is incomplete unless one efficient cause could be assigned to all observed phenomena. The concern of science, by contrast, is not to account for the productive cause of phenomena, but to discern, under what theoretical or empirical grounds, certain phenomena occur. In that way, scientists do not care for the explanation of phenomena in terms of sources. What they want is to understand those regular patterns of behavior that can be detected within phenomena, and the phenomenon is and has always been for scientists, the beginning and the end of an explanation. Nevertheless, metaphysical concerns usually are more attractive than scientific explanations, and hence, most of the time it is hard to disconnect scientific data and results from metaphysical speculations about the source of the phenomena. To understand this idea, some considerations are at stake here, which are exposed in the following section.

b) Teleological explanations

As mentioned above, when Aristotle analyzed the four meanings of the term 'cause', he wanted to explain, not four, but only three different ways of understanding changes or processes. The first was the spontaneous transitions of matter from one form to another, or from one place to another; the second was the produced transitions of matter by an external agent; and the third was our conceptual explanation of what we conceptually think is achieved with the spontaneous or produced transitions of matter. Obviously, if we consider that the third form of explanation is intellectually added when we grasp the final result of the changes observed on the other two, it is possible to assert, first, that final causes does not refer to an extra process, movement or change and, second, that for Aristotle there are, properly speaking, only two observed kinds of changes or processes: the natural or spontaneous, and those that are produced by agents, that is, by efficient causes.

Now, it is important to bear in mind that Aristotle's notion of efficient cause was introduced to explain those actions that could alter, modify or give a new shape to what he considered to be a natural or spontaneous behavior. This means that efficient causes are closely linked with our precise knowledge of natural process. Nevertheless, when Christian theologians amend Aristotle's notion of efficient causes, what they pretended to explain was not the same that Aristotle meant to explain. In order to put God as the first cause, they have to deny the idea of spontaneous processes in nature and, as a consequence, efficient causes become to be understood as the general scheme that had the purpose of explaining the source of all things whatsoever. This difference is important because while Aristotle was able to identify empirically when and how an efficient cause was operating via the observation of an altered behavior of a natural process, the Christian notion of efficient causes loosed that reference and, hence, the explanation concerning the source of both, natural and altered behaviors, becomes to be a subject of speculation. Let us try to understand why this happens.

When the conceptual structure of the Christian notion of efficient causes mix final with efficient causes, the explanatory scheme of efficient causes were altered, and the innocuous notion of final causes began to operate with the model Aristotle reserved only for efficient causes. This mixture, as we mentioned before, is the origin of the scheme that characterizes the conceptual structure of what we now call '*teleological explanations*'. This sort of explanations puts the result of the performance at the beginning of the explanation. As an outcome of this transgression, the end, that is, the purpose we conceptually assign to the performance, becomes interpreted as the efficient cause of the performance.

Three different steps are usually involve to conform the explanatory structure of a teleological explanation: the first consists in giving an accurate description of the behavior performed by a system; the second, in ascribing an end to explain that performance in term of final causes; and the third, in explaining the pattern of behavior performed by the system as an outcome of the end we previously ascribed to that performance. The latter step is rather peculiar for two reasons: first, now it seems that the system is the one which has a purpose of its own, and, second, under this conceptual scheme, the purpose has to be

interpreted as the causal source of the system's performance. As a result of this construction, we obtain that our explanation about what the system achieves with its performance, now is no longer our explanation, but the intention, desire, goal, etc., that the system had previously in its "mind" to guide its performance. An historical example might be enough to illustrate how this form of explanation operates.

During the XVI and XVII centuries, after a series of careful measurements and observations, scientists finally were able to assert that it was the earth which revolves around the sun and not the sun around the earth. They, in addition, were able to find out that the earth revolves upon its axis, and that the earth's axis had an angle of inclination. Now, if those scientist were willing to use Aristotle's form of explanation of final causes they, indeed, could be able to tell us that, as a result of the revolutions of the earth upon its axis, it is eluded the possibility that one side of the earth could burn and the other could frizz, and that the angle of inclination of the axis of the earth seems to be enough to explain the variation of seasons.

Notice that, in this example, scientists describe the phenomenon related with the earth's movements, and final causes tell us something extra: the explanation of the results that we figure out are obtained from those movements. Both explanations taken together, in this case, enhance our knowledge, but scientific descriptions and final causes do not overlap. However, let us remember that teleological explanations make a translation of final causes in terms of efficient causes and that plans, goal, desires, intentions and the like, under this form of explanation, are to be interpreted as the hidden causes of the described phenomenon. As a consequence of this procedure we find that William Gilbert (1544-1603), which is considered the father of the modern science of magnetism, explains the very same facts in the following way:

The earth perceives the sun's magnetic field, and reasoning that one side will burn while the other freezes if it does not act, it chooses to revolve upon its axis. It even chooses to revolve upon its axis at an angle in order to cause the variation of the seasons (R. Westfall, *The Construction of Modern Science*, 1977, p.27).

Gilbert combines the described behavior of the earth with the explanation of what it is obtained as a result of that behavior. When he combines both forms of explanation and translate them into the scheme of efficient causes, the structure of his "explanation" compels Gilbert to try to add something extra: the hidden (in this case psychological) causes that provoke the earth's behavior. In order to account for those causes, Gilbert has to ascribe to the earth mental faculties, such as perceiving, reasoning and deliberation, to complete his efficient scheme. He, evidently, has to assume those mental faculties as responsible for the earth's behavior, and, hence, he also has to assume that they work as the hidden factors of its behavior. Clearly, under Gilbert's interpretation, the results obtained from the earth's movements are interpreted as the effects of the earth's thoughts, intentions, goals, etc. However, it is important to notice that his form of explanation does not improve the empirical data and the understanding we previously had about the earth's behavior. Gilbert's cause-effect structure - provided by the teleological form of explanations - only induces him to assert that, in order to be able to offer a complete explanations of the earth's behavior, it is necessary to ascribe to it mental capacities.

Now, probably today we might laugh about this ridicule animism, but not necessarily because we understand the form of explanation that makes Gilbert to sustain that the earth had mental capacities. Most people tend to reject Gilbert's idea simply arguing that is totally absurd to endow the earth with intelligence. Nevertheless, when psychologists, biologists, or computer scientists allege that intelligence, or similar mental states, have to be considered as the cause of the behavior of human beings, brains or computers, what they do not realize is that they are following the very same steps that conform the conceptual scheme that characterize the teleological form of explanation Gilbert used to "understand" the earth's behavior. Unfortunately, and as we will see in the next section, Gilbert's error not only is present in the teleological forms of explanation, but in the mechanical thinking as well.

c) Occult qualities of matter and mechanical philosophy

As we have seen, when metaphysicians try to explain natural movements in terms of teleological causes, they have to sustain that the source of observable natural processes could not be explained unless we accept that matter had tendencies, attractions, desires, perceptions and all kind of the so called 'occult' or psychic qualities. Obviously, these metaphysicians did not distinguish mind from matter, hence, the so called 'theory of occult qualities of matter' has to be classified as monist, in the sense that it accepts only the existence of one substance: the material. Nevertheless, its form of explanation requires to introduce those psychic or occult powers to reside inside matter in order to explain the source of the behavior that, in its view, had a finality of their own. Against this animistic or vitalistic doctrine of matter the so called 'mechanical philosophy' raised.

It is not easy to try to explain in few words what mechanical philosophy exactly consists of. In part because this doctrine is full of inconsistencies, and in part because the people that work with this doctrine is used to those inconsistencies. Nevertheless, one fact is that it has to be considered as the doctrine responsible for the metaphysical distinction between mind and matter. The source of this distinction is the result of the conceptual effort made by its founders to eliminate all traces of psychic qualities from matter, but without abandoning the efficient form of explanation that characterizes the Christian thought. Briefly, the enterprise of mechanical philosophers was not, strictly speaking, directed toward the explanation of a new domain of phenomena, but to offer an alternative way to understand the nature of matter, and the source of its behavior.

Against the doctrine of occult qualities, mechanical philosophers sustained that matter had to be considered indifferent to its changes and movements. Nevertheless, they have to confront the problem related to the source of those movements and changes that we are able to observe in nature. This problem could be avoided taking another track of thought: to accept, as Newton later did, that one thing was to speculate about the nature of matter to explain the source of its activities, and another to try to understand its spontaneous movements without attempting to go any further. But Newton's option implied for mechanical philosophers to abandon the efficient form of explanation, something that they were reluctant to accept. What they wanted to maintain was that matter, by its nature, did not

move or change by itself, and that in order for it to move or change, an external force, agent or cause has to intervene. Unfortunately, in order to uphold that thesis, they have to confront two closely related problems. The first was to find a plausible way to take out of matter its psychic characteristics, the second was to be able to explain the source of its movements preserving intact the cause-effect scheme. Both problems were confronted in metaphysical grounds. They began its metaphysical construction asserting, first, that God created matter as a homogeneous extension with no more qualities than those that could be treated geometrically, and second, that it was God that set matter in motion. In this way, the creation of matter and the source of its motion apparently were "granted" with God (a metaphysical cause). Nevertheless, the "solution" concerning the qualities of matter gives rise to a new and unexpected problem: the one that introduce the mind-body distinction. Let us try to follow the source of this distinction.

The term 'matter' in this doctrine was not introduced as a collective noun or as a conceptual definition of the material substance in geometrical terms. This mathematical notion of matter was considered as an ontological entity, and as such, it has to be taken as the stuff that conforms all the material objects of the universe. This idea confronts mechanical philosophers with a discrepancy between the characteristics possessed by material objects and the objects we perceive, in the sense that, when we perceive everyday objects, we perceive them with much more properties that just the geometrical ones. Now, instead of going back and reconsider the fundamentals that create the inconsistency between matter and perceived objects, mechanical philosophers decide to add more metaphysical thesis.

To solve the discrepancies between the objects we perceive, adorned with secondary qualities, and the mathematical matter, for example, this doctrine introduced the idea that we do not perceive matter directly. According with mechanical philosophers we only perceive the effects that the movements of matter generate in us through our sense organs and our brain (Lara, N., Newton: Percepción y Explicación Científica, 1993). For them, the movements produced in our sense organs and our brain give rise to our mental perceptions of objects that pose colors, taste, texture, sound, smell and the rest of the properties that usually have been called 'secondary qualities'. In this way, mechanical philosophers, in order to get rid of the psychic powers of matter, had to put both, psychic and secondary qualities, inside of our minds, a new substance that has to be introduced as totally different from matter. Once this is recognized, the mind-body problem inevitable arises, combined with a series of metaphysical consequences which, even now, affect psychological and biological explanations. To illustrate this point, here, we mention briefly only some of these consequences.

Granted the existence of two different substances (matter and mind), mechanical philosophers were in the position of asserting that matter is deprived of psychic and secondary qualities. According to them, material objects produce in us not only the sensations of secondary qualities, but of different feelings, thought, etc. This idea is the source of the so called '*causal theory of perception*', a theory that is implausible but, at the same time, impossible to be refuted once we accept its most basic principles (e.g., the distinction of the two substances, the distinction between the external world and its

perception, the distinction between primary and secondary qualities). Among the consequence of accepting this theory, we find the problem of solipsism, the problem concerning the existence of other minds, the problem concerning the knowledge of the external world, the problem concerning the relation between language and thoughts, and many more.

Nevertheless, once the causal theory of perception is accepted, it is possible to establish that people tend to mistake the objects perceived through the senses with the material objects because they only have acquaintance of their perceptions. This "confusion", in their terms, is responsible for our believe of matter having color, taste, odor, etc., which, for them, it is a mistake. In the same way, they also sustain that we project upon matter our feelings, emotions, desires, etc. Matter, however, has no psychic powers, and it operates mechanically, that is, without any sensation, plan, goal, intention, etc. The contrast between teleological and mechanical explanations is settle here, in the sense that, the first is reserved to explain the human sensations, feelings, desires, etc., whereas the second is concerned with explaining the behavior of the senseless matter. According to this, the division between psychological and physical explanations is established for the first time. In principle, we must accept that the study of minds and of bodies are two totally different fields of knowledge, and that each field does not only requires a different forms of explanation but that one is irreducible to the other. The first explains the source of behavior in terms of motivations, goals, plans, etc., while the second explains the same behavior in terms of purely physical movements.

Now, mechanical philosophers assert that anything corporeal is made of particles that only posses geometrical properties, and that the mechanical movements and arrangements of those particles constitute the whole physical or external world. This idea entails the epistemic assumption known as 'determinism'. Determinism sustains that, given a complete description of the position, interactions, and movements (velocities included) of all particles at one instant of time, any event, past, present and future, can be predicted. This is, properly speaking, what constitute the main thesis of the mechanistic point of view of the world (Camap, *An Introduction to the philosophy of Science*, p. 217). The acceptance of that thesis somehow implies to deny, if not the cancellation, of all psychological forms of explanation from physics.

Nevertheless, within this doctrine we find two different approaches in the mechanical form of explanation. The first pretends to explain the nature and behavior of any physical thing only in terms of different arrangements and movements of particles. The second, goes further and explain how a world, made only of physical particles, is able to produce in us different sensations, feelings, thoughts, hopes, fears, memories, expectations, moods, humor, acts of deliberating, judging, choosing, motives, intentions, and so on.

Now, let us recall that mechanical philosophers introduce the mind-body problem, the causal theory of perception, the determinism thesis, etc., in order to avoid the tendency to explain the source of bodily movements through occult qualities. However, as part of their program, they seek to explain the source of the context of the mind through the activity of the body.

Somehow, in order to be coherent with their determinism thesis, they sustained that it is the state of the body which determines the state of the mind. The problem here is that both, the mechanical and the teleological forms of explanation preserve the cause-effect relation, and when we combine this relation with the mind-body distinction, it seems to be more logical to ascribe to the mind, and not the body, the capacities of choosing, thinking, deciding, deliberating, to judging, etc.

To solve the inconsistencies derived from the set of metaphysical thesis introduced by mechanical philosophers, there are three different alternatives:

- a) The first consist in denying the existence of mental substances and accepting a strong form of materialism. This entails to assume that all kinds of behavior could be explained in terms of the effects of its physical conditions. Strong materialism, in this way, somehow assures the possibility of taking out of matter any trace of mental capacities, but it also entails the impossibility to account for perception, knowledge and all those mental activities that mechanical philosophers themselves decided to put into the mind in order to avoid the presence of occult qualities in matter. Nevertheless, the assumption of this kind of materialism has to be seen as the doctrine responsible for the attempt to reduce the study of biological systems to physics.
- b) The source of the second alternative may be conceived as a way to solve the incompetence of this materialistic position to account for mental capacities. This alternative introduces, again in metaphysical grounds, a new form of dualism. It distinguishes biological matter from the rest of matter, and decides to put inside the biological matter the occult or psychic qualities that originally made mechanical philosophers to propose the most complicated metaphysical thesis ever constructed. This alternative is responsible of different vitalistic positions (*e.g.*, naive vitalism, organismic biology, and emergentism) that give rise to the attempt of reducing the study of the mind to the study of biological bodies. Nevertheless, well understood any vitalistic position is nothing more than an amended version of animism.
- c) The third alternative is a strong way of rejecting the step back to the doctrine of occult or hidden qualities of matter that is implied in any form of vitalism. Trying to be coherent with some of the postulates introduced by mechanical philosophy, this alternative assumed the mind-body dualism, and tried to offer a plausible way to solve the puzzle about their form of interaction. However, the impossibility to give a satisfactory answer for that interaction, can be seen as responsible to give rise to the study of mind as an entity entirely different from the body. Obviously, it is hard to understand what kind of entity the mind should be, but it originated the attempt to split biological from psychological studies.

Clearly, the problems between physics, biology and psychology can be seen as a result of these three alternatives. All three, in time, have contributed to install biological explanations in a very peculiar position, in the sense that biological studies seem to be located right in the middle of two different forms of explanation: the *mechanical* and the *teleological*. We

already talk about the steps that conform the teleological form of explanation. In the next section we discuss some aspects of the mechanical form of explanation.

d) Mechanical explanations

According with Aristotle efficient causes are only one of the many valid forms of explanation, that can be used to understand those changes and movements we observe in nature. Also, it is important to bear in mind that the specific application of Aristotle's notion of efficient cause has as its reference our previous knowledge of what he calls a 'natural' or 'spontaneous' process. In this way, for Aristotle it is clear that when we observe an alteration, interruption, modification, etc., of a natural process, we have to search for the agent that produces it. However, the Christian notion of efficient cause tries to explain, besides the alteration, interruption, modification, etc., of a natural process, the production of natural processes themselves. These very last attempt is the one that has to be considered as the source of the metaphysical speculations that give rise to both, teleological and mechanical forms of explanation.

As mentioned before, both forms of explanation seek to define what could be the causal source of all phenomena (natural and produced), and they apply, for that purpose, the conceptual scheme of efficient causes. However, for the teleological form of explanation all causal sources have to be interpreted in terms of mental items (e.g., desires, plans, intentions, etc.), which is not the case for mechanical philosophers. They take good care in dividing the realm of the mind from the realm of matter. They did not disapprove the use of teleological explanations to account for the first realm, but they did not accept the idea that the same form of explanation could be appropriate to explain the behavior of both realms.

Now, we have to recall that the main concern of mechanical philosophers was to find an alternative way to explain the behavior of matter without the intermission of mental causes as its source. Although, they were not willing to accept the idea of a natural process and, hence, they used, as the only valid form of explanation, the efficient cause. However, they did not use the combination of final with efficient causes to explain the behavior of things, which characterizes teleological explanations. Instead, they combined our knowledge of material causes with efficient causes and, following the deterministic thesis, they simply assumed that if they were able to offer a complete description of the elements and interactions that conform certain systems, they could be in good position to explain the causal source of its behavior.

Some conceptual and methodological aspects must be elucidated in order to grasp the complex conceptual procedure of the mechanical form of explanation. This form of explanation departs from the idea that the behavior normally observed in things must have an efficient cause as its source. From this point of view, they simply assume that the aim of their research is to account for that source. Using both, scientific and metaphysical definitions of matter (see Lara, N., *Emergence: A tension between definitional and empirical properties*, 1995), the most important idea of mechanical explanations is that the causal source of a system's global behavior is completely explainable in terms of the different

movements of its parts. Two different propositions are mixed in this idea. One is related with the scientific possibility of describing an observable behavior of a system in terms of the movements of its different component parts; the other is related with the metaphysical assumption that the accurate description of those movements can work in the explanation as the definition of the causal source of the system's behavior. Some conceptual confusions are at stake in this last assumption which must be clarified in order to understand the conflicts between teleological and mechanical explanations.

In general, we use the term 'behavior' to refer either to actions or to movements. Nevertheless, actions and movements are not synonymous terms. The difference between this two terms can be illustrated through an example. An arm has the natural possibility to perform certain movements. With those possible movements a system might be able to learn to perform certain actions, that is, certain possible movements which might help us to reach objects, to move other objects, etc. In this case, the difference between performing a specific spontaneous movement with the arm and performing that specific movement in order to do something with it, makes the difference between movements and actions. Actions usually are learned with experience. They are not spontaneous in the same sense movements are. Nevertheless, once we learn what is possible to obtain with certain movement or group of movements, we also learn to perform with those movements the actions that are capable to give us certain expected results. Notice that when we talk about actions we are using final causes, and the use of this kind of causes is what let us explain the difference between a movement and an action. The description of a movement by itself has no final causes, the action, on the contrary, has to have a final cause, that is, an expected (previously known) purpose, function, or direction that we can observe and understand as something conceptually added to the performance of the movement itself.

Now, it is obvious that there is nothing wrong with using final causes to explain and to understand the behavior of a system. As a matter of fact, biologist, engineers, psychologists, and social scientists, among others, will find useless to make a mechanical description of the system's movements through the analysis of their components and interactions, without knowing the function, goal, or task of their system's behavior. Unfortunately, the advantages and richness of mechanical descriptions has not been well understood, mostly because it has been used to assert that the mechanical description of the elements and interactions that the system needs to perform a specific movement, not only is the mechanical description of the system's action, but the most reliable way to try to reduce teleological forms of explanation to mechanical explanations. This belief arises simply because once we know the purpose or result of a specific behavior, it is always possible to describe the movements of the parts as if each one of them has the "knowledge" of what to do to fulfill the purpose of the action (Rosen, 1985, klir, 1991, Kampis, 1991).

However, if we stress the distinction between the mechanical description and the mechanical form of explanation, and understand that mechanical explanations use the mechanical descriptions to explain the source of both, the movements and the actions, we find that within mechanical explanations coexist three attitudes: one conceptual, other scientific, and

another metaphysical. The conceptual aspect is related with our understanding of the end of certain movement or group of movements; the scientific aspect with the analysis of one movement performed by a system in terms of its different parts and its different minute movements that makes the general movement possible; and the metaphysical with the idea that the description of the parts and the interactions associated with the global movement is the explanation of the causal source of the global movement performed by the system. Obviously, the analysis of the system is susceptible to give rise to different levels of mechanical descriptions (e.g., physical, chemical, electrical, etc.), but that description is not the explanation of the causal source of movements. Here, it seems to be necessary to remember one simple fact in order to grasp the difference between the scientific and the metaphysical statement.

It is more than obvious that in order to be able to move an arm it is necessary to have one. In the same line of thought, we may say that in order to move the arm, it must have all necessary mechanisms that permit us to move it. Thanks to science we all know that if some of those mechanisms are not working the way they normally do, it is understandable that the arm probably will not perform those movements that in the common circumstances we expect it to be able to perform. Nevertheless, to accept that the normal performance of the mechanisms of the arm, at any level of description, are necessary to move the arm, is not equivalent to say that the causal source of the movement of the arm are its mechanisms. Unfortunately, this causal statement is the one that transform scientific descriptions into metaphysical explanations. To understand the roots of this misunderstanding we have to consider other metaphysical problems usually implicated in the mechanical form of explanation.

Generally, those who employ the mechanical form of explanation tend to introduce in the mechanical descriptions two foreign conceptual mistakes. The first is the tendency to think that through the scientific description of the mechanisms it is possible to explain both, actions and movements; and the second is that when they are able to describe in scientific grounds, let us say, the movement of an arm in terms of the movements of its muscles, tendons, nerves, brain activity, etc., they tend to believe that the accurate description of those mechanisms have to be understood as the "discovery" of the causal source of the action performed by the arm.

Obviously, mechanical descriptions and teleological explanations are mixed here, and as a result of this misleading mixture, it is common to find researchers questioning themselves about how is it possibility for pure physical elements and interactions, deprived by definition of mental capacities, to "decide" the sequence of movements and interactions that the system "needs" to perform in order to obtain the "desired" (resulting) action. What those researchers do not understand is that questions of this sort arise only when we put our knowledge and interpretations of the expected action result as something previously and internally (mentally) determined by the system or its component parts. Unfortunately, as mentioned above, this step is the origin of the so called 'mind-body' problem. In recent times, there are people that believe that in order to avoid this problem, it is enough to simply declare themselves materialists. What they do not realize is that with this label they

are solving nothing, rather they are taking a step back to the all doctrine of occult qualities. With this doctrine behind the interpretations of their scientific data and results, they think there is no problem asserting that chemical elements, molecules, enzymes, genes, neurons, brains, computers, etc., are things capable to think, decide, or plan their actions.

Nevertheless, when someone remembers that the aim of teleological and mechanical explanations was to account for the causal source of everything, and ingenuously ask for the causal source of those mental capacities ascribed to matter, the so called 'materialists' try to solve this puzzle arguing either, that those occult (mental) qualities have to be understood as emergent properties of matter, or they claim that we do not know yet all the complexity of the mechanisms under study. Both excuses allow people to leave aside the healthy effort of periodically reconsidering those assumptions that may give rise to all kind of problematic interpretations with answers or pseudo-hypothesis that are more akin to bad metaphysical fictions than to science.

Now, if we detect metaphysical fictions in our explanations and are able to take them aside of our interpretations, nothing really happens with our scientific data and results. Nevertheless, the case is different when it is our own form of explanation the one that makes those metaphysical fictions necessary to appear in our interpretations. In this case, we are in trouble, most of all because now it becomes essential to revise, not our data and results, but the very conceptual foundations that are giving rise to questions that make us look for elements or mechanisms that do not exist. Think in Gilbert's example. His scientific data and results were accurate, but his interpretation was not. His form of explanation take him to endow matter with psychic powers, and the only way to avoid those psychic powers consisted in the revision of his own aim: to explain the causal source of the earth's behavior. Now, let us think in the case of those who take, let us say, life, intelligence, consciousness, etc., as properties or characteristics of some systems. If this people use the mechanical form of explanation, they will be impelled to try to find in the system those elements or mechanisms that could explain the causal source, not only of the thing itself, but of the behavior that make us call them alive, conscious, or intelligent.

What they probably do not notice is the fact that the analysis of the system will never reveal the causal source of the system, neither the elements or mechanisms that makes a system to be alive, conscious, or intelligent, because of one simple reason: those properties or characteristic are not physical (neither mental in the substantial sense). They are, properly speaking, only the adjectives we normally use to classify certain kinds of behavior. Unfortunately, when people mistake adjectives with substantives, and take those adjectives to be part of the components of a systems that display certain behavior, then, they tend to classify them as alive, conscious, or intelligent, and try to find in those systems the physical or mental elements, or mechanisms, that give them those particular properties or characteristics. The trouble here is that there are no elements or mechanisms capable to explain the nature or source of those properties or characteristics. The reason is obvious: the reference of those "properties" or "characteristics" belong to the description of the behavior, not to something that the system contains. Nevertheless, those who employ the mechanical form of explanation will always think that their research is incomplete until they could be

able to discover the mechanisms of those properties or characteristics that are impossible to find. An example might illustrate what is at stake here.

In the famous article called "*Behavior, purpose and teleology*" (1943), we find three eminent researchers, Rosenblueth, Wiener and Bigelow, looking for the mechanism of the brain that, according to them, could be the source of voluntary acts. These authors assert the following:

The basis of the concept of purpose is the awareness of "voluntary activity". Now, the purpose of voluntary acts is not a matter of arbitrary interpretation but a physiological fact. When we perform a voluntary action what we select voluntarily is a specific purpose, not a specific movement. Thus, if we decide to take a glass containing water and carry it to our mouth we do not command certain muscles to contract to a certain degree and in a certain sequence; we merely trip the purpose and the reaction follows automatically. Indeed, experimental physiology has so far been largely incapable of explaining the mechanism of voluntary activity. We submit that this failure is due to the fact that when an experimenter stimulates the motor regions of the cerebral cortex he does not duplicate a voluntary reaction; he trips efferent, "output" pathways, but does not trip a purpose, as is done voluntarily (Reprinted from *Philosophy of Science*, Vol. 10, No. 1, January, 1943, p.19).

In order to explain how the arm moves to reach a glass of water and take it to our mouth, these authors offer two different kinds of explanations: the first metaphysical, and the second physical or, as they put it, physiological. In their view, the most important one is the first, not the second, because what they want is not the physiological description of how the muscles of the arm are contracted to a certain degree and in a certain sequence, neither the description of the activity of the motor regions of the cerebral cortex connected with the movements of those muscles that have to be contracted to perform the action. What they want is the explanation of the mental causes that "trip the purpose" of voluntary acts.

Now, under their interpretation what they are calling a 'voluntary act' also has two references: the physical movements that fulfill the action, and the mental state that "trip the purpose". In the light of this interpretation there are two different kinds of actions: the physical one, which in this case is the movement of the arm that reaches the glass and take it to the mouth, and the mental action, which according with them is the volition that trips the purpose before the movement of the arm takes place. What they do not have clear is that, even if we accept the existence of the mental action as the source of the physical action, the success or failure of the physical action does not depend on the "volition" or "purpose" of the mental act. It depends on the final result of the physical movement, because the reference of a purpose is not a hidden mental state, but the result that is expected to be obtained when we when certain movements (or changes) are performed (or take place). As a matter of fact, it is dubious that we have to choose a purpose voluntarily before performing the movement of the arm to reach a glass of water. Normally once we learn how to move the arm to reach a glass of water, we simply do so automatically, as these authors somehow recognize. However, if the movement of the arm succeed in reaching the glass of water and someone ask us if the movement was voluntary or not, we might be inclined to say that the act was 'voluntary', not for describing a mental act, but to express that we performed that movement in order to reach the glass and that our performance succeeded. Notice that the

term 'voluntary' is applied afterwards, not before the movement of the arm succeeded in reaching the glass of water. Yet, if the movement of the arm instead of reaching the glass throws it, two things happen: the first is that we will explain the failure of the movement saying that the final result was involuntary; the second is that the involuntary act of throwing the glass does not invalidate the purpose of the movement, it just compares the obtained result with the expected one. This only means that both, voluntary and involuntary acts do not depend on mental decisions, choices, desires, etc., but on the result we obtain at the end of our performance.

Nevertheless, voluntary (mental) acts are introduced by these authors because, under their explanation, they need a purpose as the cause of the performance. The problem is that if they accept that voluntary acts can work in the explanation as causal sources, they have to accept that involuntary acts also can work in the explanation as causal sources. This happens because they distinguish between the voluntary selection of the purpose (the mental act) and the physical act that, according to them, follows from that selection. It is not clear how these authors could be able to explain the failure of the mental purpose to keep the causal link between mental and physical actions. However, our point here is that they need to introduce mental acts because they believe there must be a purpose behind those performances that succeed in accomplishing a particular task. Those purposes, however, add nothing to the explanation of the behavioral performance.

Unfortunately, when people takes a purpose to be a mental state, and interprets it as the causal source of the behavior, they are committed to introduce not only more mental states to explain the source of the purpose (e.g., volition, selection, etc.) but also an agent capable to perform those mental states before the behavior is carried out. That is why in the conceptual construction of Rosenblueth, Wiener and Bigelow a metaphysical agent has to be introduced to select or to choose a purpose "voluntarily". Obviously, they were not able to specify who and where that agent is, but they needed that agent to support what they were trying to say. Yet, one thing is clear, their agent has to have mental capacities to decide, to choose or to select the purpose. Therefore, they have to employ a metaphysical doctrine to account for the relation between mental and physical acts. The important result is that, in one way or another, the mind-body dilemma is embedded on their interpretation and, unavoidable, it has to be confronted.

On the other hand, probably trying to avoid the acceptance of the mind-body distinction and of mental causes as the source of the performance, these authors assert that "the purpose of voluntary acts", at least in the case of human beings, "is not a matter of arbitrary interpretation but a physiological fact". From their paper is hard to understand in what sense a purpose could be considered as a physiological fact. It is enough to remember that everything we can talk about falls into certain broad classes or categories. These categories allow us to elucidate when a statement is true, false, or meaningless. In this case, saying that purposes are physiological facts seems to be as meaningless as saying that Napoleon is a prime number. In both cases there is a category-mistake involve (Hospers, p. 88). However, if we assume that these eminent authors are using the term 'purpose', not in its normal meaning, but to refer to some physiological fact, what we could expect from them is the

description of that physiological fact to understand their reference to the term 'purpose'. They acknowledge, however, that there are no empirical basis for such physiological fact, but they must postulate its existence to support their own statement.

Despite this, what seems to be clear is that they are assuming two things: first, that even when physiologists have been incapable to find and to explain the physiological mechanism of voluntary activity, that mechanism should exist; and, second, that one of the aims of physiologists is to seek for and to explain that mechanism. What they do not realize is that the mechanism they are looking for is exactly what we are calling a 'metaphysical fiction'. The problem with these fictions is that they are not harmless, they contribute to distract the efforts and resources available for the physiologist, and, also, to convince them that somewhere in the brain, neurons, chemical elements, atoms, particles, or their interactions there must be the physiological elements or mechanisms that give rise to voluntary acts.

This is a problem because, given the distinction between movement and voluntary activity, and the assumption that both are physiological facts, this authors suggest that it does not matter if physiologists have been able to find all the physiological elements and their interactions, needed to describe a movement or the sequence of movements that are performed when we reach a glass and take it to our mouth. Neither it matters that the information derived from that knowledge is the relevant one to explain not only the success, but some mechanical failures to perform that or similar movements. They do not want the explanation of movements, neither the physiological description of those movements, which is all that physiologists need and are going to be able to find. What matters for them is to seek for the explanation of the physiological mechanisms that, according to them, give rise to mental acts, a topic which is far distant from the scientific research. Let us try to understand why.

Rosenblueth, Wiener and Bigelow not only added mental causes to the explanation of the final result of the movement people perform when reaching a glass, they also committed themselves to the idea that those mental causes have to have physical causes. In their view, hence, one task of an experimenter is to find the physical causes of mental causes. If we recall Gilbert's example, what we had was an earth having mental faculties, which, according with Gilbert, are postulated to explain the capacities of the earth to "perceive" and "reasoning" to choose the proper movements that are needed to avoid that one of its sides could burn and the other frizz. Following Rosenblueth, Wiener and Bigelow, it will be reasonable for them to ask Gilbert for the explanation of the physical mechanisms that trips the earth's "purpose".

It is obvious that Rosenblueth, Wietner and Bigelow are not talking about the earth, but about the brain, or something related with the physiology of the organism. Nevertheless, in both cases, we find the idea that in order to explain the final result of an activity as a cause-effect relationship, mental faculties have to be considered within the physical objects. In the case of Gilbert, the introduction of those mental faculties were enough to explain the earth's behavior, but in the case of Rosenblueth, Wiener and Bigelow something else was needed: the explanation of the mechanisms that give rise to those mental faculties. The result is that

above one metaphysical fiction (the mental faculties of objects) another metaphysical fiction is added: there are mechanisms that trip those mental faculties. Well understood this might sound absurd, nevertheless, we still find people trying to locate the physiological, or physical, mechanisms that cause things to be alive, to perceive, to think, to decide, etc.

Behind this kind of research we find at least five different assumptions to justify the idea that there has to be physical mechanisms to account for processes that strictly speaking the very theory that postulate their existence does not consider them physical.

- a) The first relates to the idea that behind all physical object there is a substance that gives them not only their particular shape (their formal cause) but their particular mode of behavior in certain ways (their material and efficient causes).
- b) The second is related to the belief that supports that the world we perceive is made of matter, assumption that we is responsible for the introduction of the mind-body distinction.
- c) The third relates to the idea that the observable behavior of all bodies has to be understood, and explained, in terms of the result of certain kinds of processes, states, or movements that belong to the substances that conform them.
- d) The fourth is related to the idea that in order to explain the so called 'purposive' or 'teleological' behavior, it is necessary to consider two different kinds of processes: the physical and the mental.
- e) Finally, the fifth is related with the idea that, maybe in time, scientists will be able to explain the so called 'mental processes' in terms of material causes.

Unfortunately, this last pretension takes us nowhere, in the first place because, as Aristotle well recognized, the description of an object in terms of its material causes tell us nothing about its efficiency, and second, because the only way to explain teleological processes in terms of material causes is accepting the theory of occult qualities in matter, a theory that was exactly what mechanical philosophers wanted to discard.

As we mentioned before, Newton and his followers were probably the first to recognize that neither the theory of occult qualities, nor the search for material causes could explain us something about how in fact things spontaneously behave. That is why, for him as well as for others, it was more than enough to study, in theoretical or observable grounds, the regularity of natural or experimental processes to obtain a complete scientific account of the phenomena under study. Nevertheless, neither in Newton's time, nor in ours, people has been able to avoid the temptation of trying to add a metaphysical explanations concerning the causal source of natural phenomena. To illustrate this idea, in the next section we introduce ourselves in our last form of explanation: the physical one.

5. Physical explanations

For its mathematical treatment of matter, people easily tend to confound mechanical philosophy with the Newtonian mechanics. However, mechanical philosophers conceive matter as a substance, whereas Newton conceive it as a definition. This difference is

important because when we think of matter as a definition we are not saying that things are made of matter, but only that we are stipulating what properties, or characteristics, are enough to rightly call something 'material'. This means that while mechanical philosophers were impelled to consider matter as an extra thing behind the perceived objects, for Newton the term 'matter' was simply a concept to defined what he was going to consider a physical body, not the extra object, or substance that conform the perceived world. Taking in consideration this distinction, we may say that Newton's mechanics do not have to face the problem concerning the difference between perceived and physical objects, a problem that is the source of the mind-body distinction. Neither he had to accept the ontological compromise of sustaining that things are made of matter, which is the origin of the doctrine call 'materialism'. As a matter of fact, Newton never pretended to reduce the study of everything to physics, but only to define the most basic principles and concepts to derive the laws of movement of what he defined as a physical body.

In this respect, it may be said that while the concern of mechanical philosophers was concentrated in finding out the elements or interactions that, according with their expectations, were necessary to explain the causal source of a system's behavior, the concern of Newton's mechanics is concentrated in finding out if all the solids, liquids and gases that is possible to find in bodies (whether alive or not), share the same basic properties which define a physical body, and if all of them naturally follow the very same rules or laws of movements when no external agent intervene. Understanding this contrast, we may say that Newton, in a sense, is much more akin to Aristotle than to mechanical philosophers. Nevertheless, the resemblance between Newton and Aristotle should be taken carefully, hence, some important remarks seem to be at stake here in order to grasp the originality of the Newtonian form of explanation, particularly with that related with his idea of natural movement.

Let us start by considering that Biology predictive power usually requires the knowledge of the regular patterns of behavior that favor a system to reach certain ends. By contrast, the predictive power of physical description lays not on the knowledge of the finality of the system, neither in the component elements of the system, but rather on the knowledge of its movements, whether they have an identified purpose or not. Based on this reflection, we would like to point out that the difference between biological and physical interests lays in the fact that the study of the first is closely related to the understanding of actions, whereas the study of the second is closely linked to the understanding of movements. Obviously, the knowledge of movements always enhance our comprehension of actions, and the knowledge of movements not only is independent of our knowledge of actions, but it is more basic and general than our knowledge of actions.

Aristotle was more a biologist than a physicist, thus, he was more concerned with the understanding of how particular things, or classes of things, spontaneously behave towards certain end, than in the analysis of those general movements they obey. For him, for example, it was very important to recognize the morphological variety between different seeds, embryos, organs etc., in order to be able to predict the final state or result of those things or classes of things he was studying. Unfortunately, the empirical procedure of his

biological spirit prevent Aristotle from looking further in the theoretical realm of generalities, which was necessary for the foundation of a good physics. Let us try to explain this idea in more detail.

The empirical observations of the movements of heavenly and terrestrials bodies, made Aristotle think that there was no relation between them. For him, it was more than clear that the natural tendency of the celestial bodies were the circular lines, while the natural tendency of terrestrial bodies were the straight lines. On the basis of his empirical observations, Aristotle concluded that there were two physics: the celestial and the terrestrial.

In order to put together both physics Newton questioned the very idea that the plain empirical observation were enough to make us recognize the difference between natural and produced movements. It could be possible that the movements we observe in the heavenly and terrestrial bodies were not as natural as Aristotle thought. With this reflection, Newton began to speculate about what movement or patterns of movements could be enough to link the behavior of both, the celestial and the terrestrial bodies. The history tell us that the fall of an apple was his inspiration, but a lot of theory has to be introduced in order to explain the relation between the fall of a terrestrial body with the observable movements of heavenly bodies. Based on Galileo's studies of free fall, Newton thought that if the fall of the apple were not stopped by the earth's surface, and given ideal conditions, the apple will spontaneously continued its trajectory at constant speed in a right line. But once the apple is stopped by the earth's surface, it will continued in its state of rest until something moves it. This reflection gives rise to the principle of inertia, a principle that we have to recognize was not conceived observing the movements of bodies, but form idealizing condition incapable of being realized in fact. Nevertheless, these conceptual principle has to be considered the cornerstone that helped Newton to derive not only his three laws of motion, but the frame of reference that serve him as the point of departure to link celestial and terrestrial movements, as well as to understand those possible movements that can and cannot be perform by bodies, whether alive or not.

This principle asserts that bodies, by definition, do not change spontaneously their state of rest or uniform motion in a right line. Behind this principle Newton is sustaining two important things: the first is that motion and rest do not require a cause or a force to be explained, he considered both the natural tendencies of bodies; the second is that when we observe that a body is not in its inertial state, it is possible to deduce (or in his terms, infer, in theoretical grounds), that the body in question has been affected by a force (Newton, *Principia*, definition III, p.2) notwithstanding how that force is going to be defined.

Through the principle of inertia Newton put natural movements, efficient causes and forces as part of a theory, meaning that the conceptual structure of the physical form of explanation does not follow the Aristotelian tradition of the four causes, neither the Christian tradition that reduces all forms of explanation to the cause-effect relationship. The physical explanations take the form of an if-then relation, in the sense that only if we accept what the principle asserts, we may conclude whatever is derived form that principle.

In order to avoid confusions, it is important to keep in mind the difference between a conditional statement that takes the form of an if-then relation and those statements that take the form of a cause-effect relationship. The first establishes the relation that holds between technical concepts (*i.e.*, definitions that fix in advance the limits of the meaningful use of a concept), and objects or facts. The second is a temporal relationship that holds between particular object or events, where one of them is said to be the source of the other. Here we present an example to clarify the difference between these two kinds of statements.

In the case of the principle of inertia, Newton asserted that any body that is not in a state of rest or of uniform motion in a right line, has to be considered as not being in its inertial state. Notice that Newton did not mean that at the beginning of time every body was in its inertial state, because saying that would introduce him in a temporal relation about the state of the world in some time, and the change of state of the world in another time. Not only that, but also Newton get the compromise of using the cause-effect relationship in order to be able to explain the source of that change of state, something that he clearly wanted to avoid. As mentioned before, cause-effect relations are always temporal relation between objects, facts or states. But the principle of inertia does not establish a temporal relation between object, facts or states. Newton is talked about any object, and he established relations between his theoretical principle and the observable behavior of object. This last remark is important because the Newtonian form of explanation does not deal with particular things or events, a restriction that conforms one of the most important characteristics of the cause-effect relationship. It only deals with the behavior of things of events, whatever this things or events are. Newton, first, just gave us the definition of the concept, or principle to identify when something is, or is not, in its inertial state. Given this reference the if-then relation could be established between the defined behavior and the one being observed.

It is clear when that with terms like 'inertia' and 'force', Newton only asserted that if, for example, the motion or trajectory of the moon and the planets were different from a straight line, it was possible to "infer" (deduce from his definition) that those bodies were not following the pattern of behavior the principle of inertia states. Only in that sense, it is possible to assert that the observed behavior of those bodies are affected by a force. He called the force that deviates planets from its inertial state 'gravity'. However, notice again, that gravity was not conceived by him as an efficient cause (*i.e.*, a power, or an active agent), but only as the appropriate inferred postulate to explain and predict the motion of the moon and the planets. He established, in theoretical and empirical grounds, some constant relations between the inferred force of gravity and the planetary movements. But gravity was not postulated to explain the source of the planetary movement, only the relation that holds between a theoretical definition (inertia), the observed planetary movements, and the inferred force (gravity) that permit us to calculate the amount of deviation of the inertial movements of those bodies. Nothing more is implied in this form of explanation.

The if-then relations does not deal with what is behind of what can be observed, as the cause-effect relation has to do. It deals with the possibility to explain and predict how, in fact, things behave. However, the problem of understanding the structure of physical explanations is that it is always possible to translate them into a cause-effect relationship.

Mechanical philosophers, for example, conceived gravity as the explanation of the source of the planetary movements. This might seem innocuous, but when we understand gravity as the source of the planetary movement, gravity loses its status of explanatory force and becomes an ontological agent. When doing this, it does not matter if the concept of gravity help us to explain, measure and predict many observable movements. Metaphysicians would like to go even further and require the explanation of the nature, place and ways in which gravity operates on bodies to produce the observed phenomena. In this case, Leibniz (N. Lara, *Materia, Newton y Espacio Vacío*, 1992), among others, wanted Newton to explain if gravity was conceived by him as an occult quality of matter, or as the effect of a non yet discovered mechanism that pull or push the inertial trajectory of those bodies. Newton's answer to that request was the following:

... I could not discover the cause of these properties of it [gravity] from phenomena. For I flee from hypotheses whether mechanical or of occult qualities. They are harmful and do not engender science. And it is enough that gravity should really be granted and act according to the laws which we have explained, and should suffice to account for all the motions of the celestial bodies and of our sea. (I.B. Cohen, *Introduction to Newton's Principia*, p. 243)

6. Conclusion

Among the different forms of explanation we have studied in this paper, we found three traditions:

- a) the theory of the four causes introduced by Aristotle;
- b) the Christian tradition that gives rise to two forms of explanation: the teleological and the mechanical; and
- c) the Newtonian tradition, which is said to represent in our time the prototype of the scientific form of explanation.

Obviously, teleological and mechanical forms of explanation go far beyond Christianity. However both forms of explanation are unavoidable linked with the Christian thought and particularly with its compulsion to try to explain the source of everything through agents (named them God, Nature, Mind, Matter, Evolution, Movements, Processes, Forces, etc.).

Based in our discussion, we may conclude that the only tradition that systematically introduces conceptual problems is the Christian tradition. Nevertheless, it is the most fashionable in our time, and the other two almost systematically tend to be interpreted at its light (see Lara, N, *Emergence: A tension between definitional and empirical properties*, 1995). This has provoked the belief that the kind of explanation science should provide has to be done based only on two alternative forms of explanation: the teleological and the mechanical (Carnap, 1966, von Wright, 1974, 1971, Hospers, 1967, Rosen, 1985, Klir, 1991). Both forms of explanation have embedded the idea that it is never enough to understand the behavior of nature in terms of its regularities. Both want to explain the source of those regularities. It has been established, on the one hand, that teleological explanations aim to account for the causal source of those regularities in terms of purposes, intentions, desires, etc., and, on the other hand, that mechanical explanations do not

represent a solution to that attempt, because they aim to account for those alien mental purposes, intentions, etc., in mechanical terms. We may conclude, then, that teleological and mechanical explanations of mental causes takes us nowhere. It is enough to comprehend that once we understand what can be obtained by doing certain movements, we should avoid the tendency to add to that behavior that end as its cause. Ends are nothing more than our predictions or expectations of what the result should be, and not a mental cause. In this sense, it is futile to search for mental causes inside nature to explain the results of its behavior, and it is not only futile but absurd to seek for the physical mechanisms underlying those mental causes, which, by the way, does not exist.

Now, the success of mechanical explanations over other kinds of explanations, particularly in biology and engineering, has been the introduction of the modality that consist in offering different levels of mechanical descriptions related with the performance of actions. These descriptions indeed help us to understand how different elements and their interactions are combined to execute a particular movement or sequence of movements, and, without doubt, the knowledge of those mechanisms are very enlightening to find, repair, control or predict some failures or deficiencies that prevent, obstruct or interrupt the normal execution of those movements. Unfortunately, the artificers of mechanical descriptions were unable to realize the richness of their own contribution. They introduce this kind of description trying to find inside the system the purpose of its actions. Looking for purpose their contribution looses its explanatory power and makes people believe that maybe in the future they will find the appropriate material causes or mechanisms to explain the source of perceptions, desires, intentions, etc. The cost of this metaphysical dream has been immense, not only for the waist of lives, resources and time, but also in terms of the impediments it imposes to help scientist from understanding what they are looking for, and, also, what exactly they really are achieving through their own research:

- a) the explanation of the mechanisms that the system needs in order to be able to perform certain movements, or
- b) the explanation of the mental and physical causal source of those movements.

Obviously, scientists are capable to look, find and explain the first, but for reasons that we have already given, they would not be able to explain the second.

On the other hand, for us it is clear that between mechanical descriptions, Aristotle's theory of the four causes and Newton's form of explanation there is no conflict of interests. Aristotle tradition seems to be the most suitable way for understanding, linking and explaining many of the complex aspects related not only with the studies of biological systems, but with all those systems that, in one way or another, require our knowledge of the final result of their behavior. We have seen that our knowledge of final causes requires the collaboration of different forms of explanation, and that the correct use of final causes does not compromise us with accepting the existence of mental causes.

In regards to the Newtonian tradition, it simply does not work with final causes, and its form of explanation does not need the collaboration of other forms of explanation. Nevertheless, it does not mean that we are, or are going, to be able to explain everything with physical

laws and concepts. As we have seen, physical laws and concepts do not tell us how different systems might, or should, behave in normal conditions. What they tell us is that whatever a system might be able to do, it cannot trespass the limits prescribed by physical laws. In this sense, it does not matter if we are talking about heavenly bodies, molecules, stones, apples, animals or human beings, we can be sure that all of them respond to the laws of movements in the very same way. In this sense, our knowledge of the physical laws can give us a good platform to understand, not what the system in fact is capable to do, but only the limits of its performance. The most fantastic enterprise that we might be able to imagine as workable, has to be supported by the knowledge of the limits imposed by physical laws. In this way, we may say that one of the most interesting characteristics of the physical form of explanation is that it does not distinguish differences between different kinds of systems. All of them are treated as mere physical bodies, and the physical laws only tell us the limits of their doing.

Nevertheless, most disciplines need to know apart of what physics demarcates, what in fact their systems do or are capable to do, under normal or experimental conditions. Furthermore, other disciplines need to know what exactly is, or is going to be, achieved by their systems at the end of their performance. In this case, physics is mute, and we require other forms of explanation.

As mentioned above, with the theory of the four causes, Aristotle provided us with three different, and perfectly compatible forms of explanation to understand and predict both, natural and produced processes of specific systems or classes of systems. Mechanical descriptions somehow can be seen as the method needed by Aristotle to understand the use and application of material causes. Material causes for Aristotle were nothing more than the description of the elements than conform a system. Mechanical descriptions enhance that notion incorporating, in addition to the knowledge of the elements that conform a system, the knowledge of their interactions. With the help of mechanical descriptions we are now in the position to explain not only produced changes in terms of efficient causes, but failures in some previously known performance in terms of lack of elements, or in terms of alterations in their interactions. With this in mind, the following example might illustrate how different forms of explanation collaborate with each other.

Think in a computer. We know what a computer is able to achieve when we are using it, that is, we know its *final cause*. Imagine, that suddenly the computer does not do what we expect it to do; let say, that we cannot turn it on. The first thing we might try to do is to find the *material cause*, that is, if a physical element is damage, broken, out of place, etc. If it is not the case, we can take our machine with an expert to check its circuits. Obviously the work of the expert consist in making a mechanical description of the machine to try to find what is wrong with it. In order to be able to make a *mechanical description* the expert will need, first, the wiring diagram of the system, that is its *formal cause*, and second, to know its dynamics, that is, its *normal performance*. Let say that after checking all its physical parts and its interactions what he finds out is that the computer's fuse is broken. If our computer is in good hands, after finding the material cause, the expert will try to find out what was the

efficient cause that broke the fuse. In this case, his knowledge of the *physical laws* can provide our expert with good clues to find it out.

With this example is easy to see how different forms of explanation can be used to find an answer to our questions. We usually may combine them in order to understand the work not only of artificial systems, but of natural systems as well. Our understanding of their normal performance, their elements, and their interactions, as well as the final result of their performance are necessary for us to be able to predict, control or repair them. The only question that we have to learn to avoid is the one related to the causal source of natural process. In this work we have tried to show that conceptual problems only arise when we introduce this kind of questions. They, indeed, are very attractive, but we have to learn from Newton that questions of these sort engender metaphysics, but no science.

Acknowledgment. I would like to thank Dr. Francisco Cervantes-Pérez, Luis Mendoza and Francisco Caviedes for their help and valuable comments on the manuscript.

BIBLIOGRAPHY

- Aristotle, *Metaphysics (Metaphysica)*, Britannica Great Books, The University of Chicago, translated by W. D. Ross, 1952.
- Aristotle, *On the Soul (De anima)*, Britannica Great Books, The University of Chicago, translated by J. A. Smith, 1952.
- Aristotle, *Physics (Physica)*, Britannica Great Books, The University of Chicago, translated by R. P. Hardie and R. K. Gaye, 1952.
- Campbell Mossner, E., *Encyclopedia of Philosophy*, Vol. 2, Macmillan Publishing Co., Inc. & The Free Press, New York, Collier Macmillan Publishers, London, 1967.
- Carnap, R. *An introduction to the Philosophy of Science*, edited by Martin Gardener, Basic Books, Inc., N. Y., 1966.
- Hepburn, R. W., *Religious Doctrine of Creations*, *Encyclopedia of Philosophy*, Vol. 2, Macmillan Publishing Co., Inc. & The Free Press, New York, Collier Macmillan Publishers, London, 1967.
- Hopfers, J., *An Introduction to philosophical Analysis*, 2dn. Ed., Prentice-Hall Inc., Englewood, Cliffs, N.J., 1967.
- Kampis, G., *Self-modifying System in Biology and Cognitive Science*, Pergamon Press, Great Britain, 1991.
- Kerferd, G. B., *Aristotle*, *Encyclopedia of Philosophy*, Vol. 1, Macmillan Publishing Co., Inc. & The Free Press, New York, Collier Macmillan Publishers, London, 1967.
- Klir, G. J., *Facets of System Science*, Plenum Press, N. Y. & London, 1991.
- Lara, N., *Emergence: A tension between definitional and empirical properties*, *Advances in Synergetics, System Research on Emergence*, Vol. II, Edited by George E. Lasker & George L. Farre, Published by The Institute for Advanced Studies in System Research and Cybernetics, USA, 1995.
- Lara, N., *Materia, Newton y espacio vacío*, *El Concepto de Materia*, Colofón, México, 1992.
- Newton, I., *Mathematical Principles*, translated into English by Andrew Motte, 1729, edited by Florian Cajori, University of California Press, Berkeley & Los Angeles, 1966.
- Rosen, R., *Anticipatory Systems, Philosophical, Mathematical & Methodological Foundations*, Vol. 1, Pergamon Press, Great Britain, 1985.
- Rosenblueth, Wiener & Bigelow, *Behavior, purpose and teleology*, Reprinted from *Philosophy of Science*, Vol. 10, No. 1, January, 1943.
- Tomasini, A., *Filosofía de la Religión*, 2^a Ed., Grupo Editorial Interlinea, México, 1996
- Von Wright, G. H., *Causality and Determinism*, Colombia University Press, N. Y. & London, 1974.
- Von Wright, G. H., *Explanation and Understanding*, Rourledge & Kegan Paul, Cornell University Press, 1971.
- Westfall, R. S., *The Construction of Modern Science*, Cambridge University Press, USA, 1977.
- Wick, Warner, *Aristotelism*, *Encyclopedia of Philosophy*, Vol. 1, Macmillan Publishing Co., Inc. & The Free Press, New York, Collier Macmillan Publishers, London, 1967.

Methodological Considerations in Cognitive Science

Nydia Lara* and Francisco Cervantes-Pérez**

**Centro de Instrumentos (PUIDE), Universidad Nacional Autónoma de México
Apartado Postal 70-106, México, D.F.C.P.-04510*

***División Académica de Computación, Instituto Tecnológico Autónomo de México
Río Hondo no. 1, México, D.F., C.P. 01000*

Abstract

This book is the result of a meeting where neuroscientists and computer scientists explored the possibility of developing common ground for the study of intelligent behavior in animals and machines. Here, we analyze some of the methodological obstacles that create conflicts among the different disciplines that form Cognitive Science. Our aim is to set forth specific methodological considerations that might lead towards the amelioration of the Isolation Problem which results from the fact that each area of Cognitive Science has been developed in almost complete independence of the others. We argue that Isolation Problem has been in great part generated by the misuse of linguistic terms associated with mental events.

1. Introduction

Cognitive Science is the name given to a group of disciplines that share a common concern with the understanding of intelligent behavior in humans, animals, and machines, under either the serial or parallel information processing paradigms. Some of the disciplines that have participated in Cognitive Science are Cognitive Psychology, Computer Science (Artificial Intelligence), Linguistics, Philosophy, and Neuroscience (Artib, 1989; Simon, 1981; Simon and Kaplan, 1990; Sillings et al., 1989). Although everybody keeps in perspective that the main goal of Cognitive Science is to understand the causes of intelligent behavior, different fields seek different kinds of causes (e.g., mechanical, linguistic, or mental) and, thus, use different research methodologies. This gives Cognitive Science the richness of a broad interdisciplinary field and, at the same time, it also creates the problem of reconciling a variety of methodologies with very different points of view with respect to:

- a) the kind of processes that are relevant and, therefore, have to be examined (e.g., a program, the dynamics of the central nervous system [CNS], or a mental process);
- b) the approach one needs to apply (e.g., computational, neurophysiological, psychological, logical, or grammatical); and
- c) the strategy that has to be followed (e.g., top-down, bottom-up, or middle-out).

This diversity of methodologies yields a situation where we have a set of areas that give different kinds of causal explanations for intelligent behavior (i.e., explanations given in computational, neuronal, or mental terms). This multiplicity of approaches, methods, proposals, and ontologies has given rise to the unavoidable tendency to generate data and knowledge in each discipline without taking into consideration the other disciplines and their respective results. Thus, each area has been developed in almost complete independence of the others, generating in this way what we have called the Isolation Problem (Cervantes-Pérez, 1989).

Another source of the Isolation Problem comes from different explanations, mainly in Cognitive Psychology, Neuroscience and in Artificial Intelligence, where people has felt the need to turn to some kind of metaphysics in order to relate mental terms with neuronal dynamics and computational programs. Unfortunately, the recourse to metaphysics to put together accepted outcomes in the different branches of Cognitive Science has led to the unexpected result of bringing into the very root of Cognitive Science the intractable philosophical question known as the mind-body problem. This problem is still an unsolved philosophical puzzle, and has been imported into Cognitive Science by the use of mental terminology without precise definitions (see below). The outcome of using undefined linguistic terms and its mixture has been to generate incredibly many problems so that, instead of unifying the different approaches to understanding the causes of intelligent behavior, it has created many disagreements which have opposed not only philosophers and scientists but also philosophers and philosophers (for a clear example see the debate between Searle, 1990, and the Churchlands, 1990), and scientists and scientists as well.

2. The Origins Of The Isolation Problem

In Cognitive Science we may recognize three main trends of research in the study of intelligent behavior: a) those that study mental processes as responsible for intelligent behavior; b) those that identify a specific coordinated program that simulates intelligent behavior; and c) those that analyze natural or artificial

mechanisms that may underlie a specific behavior. Cognitive Psychology (Wiederman, this volume) represents the first trend, Artificial Intelligence the second (Barto; Durfee; Riseman; this volume), and Neuroscience (Nicolli; Bargas; Mendell; Wurtz; this volume) and Brain Theory (Martinez-Urieegas; Arbib; Houk; Cervantes-Pérez; this volume) the third one.

In trying to build bridges between disciplines that study mental events (e.g., Philosophy and Cognitive Psychology), or simulate some aspects of what is called intelligent behavior (e.g., Cybernetics and Artificial Intelligence), or analyze the mechanisms underlying the elicitation of that kind of behavior (e.g., Neuroscience), people has used, very "loosely", mental (psychological) terminology to unify their different perspectives. For example, to link the study of mental processes with the synthesis of coordinated programs (Cognitive Psychology with Artificial Intelligence) some cognitive scientists (Fodor, 1975, 1980, 1990; Block and Fodor, 1980; Pylyshyn, 1973, 1980; Rogers, 1964) have postulated that mental processes are like computer programs produced by the mind and that later can be executed by a proper "body". They introduce metaphysical problems by assuming a dualist position, and, in addition, they neglect the importance of the study of the mechanisms subserving intelligent behavior. In this way, they deny the possibility of unifying the work of Neuroscience and Brain Theory with that of the first Cognitive Sciences (i.e., Cognitive Psychology, Artificial Intelligence and Linguistics).

Other groups, mainly in Neuroscience, have tried to bridge the gap between the study of the purposively mental processes involved in the elicitation of intelligent behavior, and the analysis of the dynamic activation of the underlying neural mechanisms (Georgopoulos, 1989; Romo, this volume). They assume that mental and mechanical processes are the same (i.e., a monist-materialistic position), and that the same signal can produce simultaneously two kinds of effects: the spatio-temporal activity presented by the mechanisms controlling the behavior, and the corresponding intentional purpose of the organism to execute it. Fortunately, adding mental terms to the explanation of purely mechanical processes does not affect the way in which we may control, manipulate, modulate and predict their spatio-temporal characteristics; however, it introduces the problem of how a mechanism (e.g., a neuron, a complete brain, or a computer) can perceive, imagine, think, plan, wish, etc. The complete description of the mechanisms and how they evolve in time does not solve the puzzle of how mental events could be carried out by a neuron, the brain, or the computer. Unnecessarily, the mind-brain problem is

roduced again, in a very complicated manner, because now not only humans can receive imagine, think, etc., but brains, neurons, computers and many other mechanisms as well.

These are just two examples of how the introduction of metaphysics blocks the finding of a global unifying methodology for Cognitive Sciences, and even crosses the scope of the Isolation Problem. Here, philosophy can play an important role in the attempt to clarify this problem: a) by making clear what its sources are; b) through an exploration of the possible connections between the descriptions of intelligent behavior in psychological language and those given in mechanical terms. Let us take an example to illustrate what is at stake.

The word "intelligence" clearly has mental connotations and, like most mental terms, it is not clear what it is and where it could be located. The same word can be used either to point to a supposedly mental process, or just to evaluate the outcome of a specific performance (Cohen and Feigenbaum, 1981; Fischler and Firschein, 1987). This ambiguity, inherent in the uses of the same term, bears with it several, not always well differentiated, kinds of causal compromises to explain the sources of intelligent behavior. In this case, mental as well as mechanical processes are possible candidates to become the causal sources of intelligent behavior. However, intelligence, seen as a mental process, not only lacks an unambiguous definition but precise location as well. Therefore, mental events can be something that is produced by a mind, a mechanical device (natural or artificial), or as something that can emerge from the organization of special mechanical devices. This ambiguity may lead us to understand mental processes as: a) the causes of intelligent behavior;

b) nothing but neuronal or mechanical processes; c) an effect of qualitative changes emerging from the activity of organized matter; or d) something that is completely irrelevant to explain intelligent behavior. In addition, this situation makes it also possible to understand mechanical processes as: a) the source of mental processes, b) the source of behavioral performance or c) the source of both.

This proliferation of what at first sight are unrelated usages of one and the same term may be seen as the cause of believing that we have at least three elements that seem to be required to account for what we call intelligent behavior: a) the performance itself (behavior), b) the mechanical processes underlying it (neural or computational), and c) the mental processes involved (i.e., purposes, intentions, thoughts, and understanding).

3. Methodological Obstacles to the Integration of Cognitive Science

Although all cognitive scientists seem to agree that the subject of study is intelligent behavior, there is much disagreement on what to study as the cause of such behavior. For example, some cognitive scientists (Fodor, 1975, 1980; Pylyshyn, 1973, 1980; Simon and Kaplan, 1990; Rogers, 1964) believe that the only thing that matters is the intelligent process that defines the behavior to be performed, regardless of the hardware that executes it; whereas others believe that the only way we can explain intelligent behavior is through the study of the dynamics of the underlying neuronal processes (Georgopoulos; Romo; Rudomin; this volume). The former group proposes to study intelligence itself (as a mental process, or as an algorithm), while the latter group studies intelligent behavior as a mechanical (physical) process. Arbib (1989) offers an intermediate position in that intelligent behavior is explained in terms of the interactions of functional units called schemas, but these schemas are seen to be implemented in terms of neural networks (see also Arbib, this volume). It must be noticed that different approaches require different, and sometimes incompatible, methodologies. In this case we have at least three, not always well differentiated, kinds of causal compromise to explain the sources of intelligent behavior: a) the mind, b) an algorithm, and c) the underlying mechanisms. All three are to be understood as "internal" causes of intelligent behavior, but they do not share the same characteristics, or the same essence, or, even less, the same way to study them.

In addition, the outcome of this proliferation of causal sources of intelligent behavior is that cognitive scientists offer rather different kinds of explanations. Depending on whether or not they accept that mental processes have to be considered in order to account for intelligent behavior, and on where and how mental processes are to be or not to be explained, we can find: a) teleological, and b) purely mechanical causal explanations.

In Cognitive Science we have something that seems to be a third type of explanation although it is only a description of behavioral patterns. But, because of its characteristics, it is easily interpreted as a teleological explanation, when it only describes behavior in purely mechanistic terms. This is explained below, but what we want to stress now is that, strictly speaking, only mechanical causal explanations count as scientific explanations. Teleological explanations are, in fact, the opposite of mechanical explanations in the sense that the use of "reasons" or "motivations" implies that there are some hidden qualities in both objects and processes which

flow the behavioral phenomena to occur. When we say that somebody turns on the heater because he is cold, we are describing this behavior in teleological terms. There seems to be a reason for turning the heater on, and the behavior appears as the result of a desire. "Reasons" and "desires" seem to refer to hidden causes, and the behavior seems to be guided by causes of this kind. In contrast, we can explain the same action in mechanical terms by saying that when a person turns on the heater, his action was in response to a reduction in the room temperature. Mechanical explanations are concerned with the measurement of related (causal) changes in the environment and the internal milieu and changes in the activation of certain mechanisms. In the latter case, then, there is the compromise of explaining the behavior in purely observational terms; whereas in the former situation, explanations are concerned with goals, intentions, aims, and purposes. Thus, it is clear that different kinds of "causes" are involved in different kinds of explanations, which implies that people who provide one sort of explanation do not even share the same subject matter of study with people providing the other kind of explanation. [Editor's Note: This very issue of relating teleology to causal mechanisms was at the heart of a precursor (half a century ago!) of the present meeting — the collaboration between Arturo Rosenblueth of Mexico and Norbert Wiener of the USA that helped lay the foundations for Cybernetics, "the study of control and communication in the animal and the machine". See Rosenblueth, Wiener, and Bigelow, 1943.]

Therefore, it is clear that cognitive scientists are facing a methodological problem, in the sense that they do not share the same kind of ontological and epistemic basis to define: a) what to study as causes of behavior, b) what is the essence of those causes, c) what their location is, and d) what approaches and techniques can be used in order to study them. This problem in Cognitive Science is more related to Philosophy of Science than to the empirical Sciences; that is, its solution requires a review of the ontological and epistemological basis of different disciplines, in order to establish whether or not there is a possible connection between those disciplines focused only in the study of intelligent processes (regardless of the hardware that executes them) and those more concerned with understanding the mechanisms underlying intelligent behavior.

4. Models Of Intelligent Behavior

In seeking to build an automaton that performs intelligent behaviors, we find two different groups of researchers: a) those trying to transform mechanical explanations of the central nervous system dynamics into artificial or theoretical models that simulate the execution of intelligent behaviors (Arbib; Barto; Zak; this volume); and b) those taking a computational approach to build a system (algorithm) that allows a computer to carry out a similar task, without being concerned with imitating the way it is done by living organisms (Lee; Marroquin; Riseman; this volume). In both cases, we find people concerned with the simulation of intelligent bodily movements, and others who aim to model rational processes as well.

Depending on different research techniques, we can find basically four kinds of models to explain, or simulate, the so called "internal" processes of "intelligent" behavior. These models are: a) biological (neuronal); b) formal (logical-mathematical); c) computational (conceptual); and d) artificial (machines).

The first ones are used to explain any sort of animal behavior (including that of humans) in terms of different levels of neural mechanisms and bodily movements. The other three seem to require some terminology that is strongly akin to the teleological one in the sense that goals and purposes seem to be required. That is, the way they elaborate models of "intelligent" behavior requires prior knowledge of the possible outcome of two different kinds of processes: one is related to the logical sequences that have to be performed; the other is related to the translation of those logical (rational) sequences into bodily movements.

At first sight, logical (reasoning) sequences belong to a different realm from neural mechanisms and bodily movements. Presumably, the former is a "mental" process, while the second is, undoubtedly, a purely physical process (Putnam, 1960).

Common sense tends to induce in us the belief that in any kind of intelligent behavior, reasoning always plays an important role in its production. This introduces mental processes as a causal source of all kinds of intelligent behavior. However, we must point out that models of rational behavior seek to explain "reason" as an outcome of a specific kind of physical performance. Hence, the meaning of "reason" within this kind of model cannot be understood as pointing to "mental processes" at all. The sources of behavior for builders of "cognitive" models of intelligent behavior are always explained in purely mechanical terms. Nevertheless, their goal is seldom the physical underlying mechanisms themselves.

They rather concentrate on the relation that can be established between different kinds of effective procedures of behavioral activities (known as algorithms) and the production of concrete behaviors based on them. The latter feature does endow these models with a certain teleological flavor because, somehow, it is accepted that the effective procedure (algorithm) already contains the right signs that are able to express (by means of different mechanical devices) the outcome of a purely physical behavior carried out by certain rules. But, only by mistake can we assert that an "effective procedure", or expressed algorithm, is the causal source of any kind of behavior. Algorithms only describe processes that can be performed, but they are not things (physical or mental) that can by themselves generate any kind of activity. If we look at algorithms as things guiding behavioral processes, it is tempting to think of them as causes of behavior. But as soon as we understand that algorithms are not things performing processes (plans, goals, intentions, etc.) but descriptions of processes that can be performed by things, the relationship between mechanical devices and formal devices loses automatically the mind-body flavor usually considered in the teleological kind of explanation.

The two steps of an algorithmic description (a logical sequence of steps and its translation into physical movements) form a redundant description of the same phenomenon. It is a common mistake to interpret the algorithm as the cause of the behavior, but we must keep in mind that an algorithm is nothing more than the description of steps that have to be followed in order to execute the proper bodily movements. Therefore, the explanation of one statement in terms of another statement does not mean that we have found the cause of the behavior to be understood. In fact, when we try to study the underlying mechanisms of the performance, we have to use a mechanical explanation. That is, an algorithmic description of behavior cannot be a teleological explanation of behavior, but mechanical.

When we understand the main characteristics of the different kinds of explanations, it is clear that it is possible to build a bridge between algorithmic descriptions (Artificial Intelligence) and mechanical explanations (Neuroscience), keeping in mind their different levels of analysis. One describes the logical sequences of the behavior; the other explains the underlying mechanisms that make possible the behavior.

5. Neuroscience, Cognitive Psychology, and Artificial Intelligence

The aim of this section is twofold: first, to review some attempts to link neuroscience to other cognitive sciences; and second, to set forth some of the characteristics that have to be considered in the development of a global unifying methodology to give proper and not redundant "causal" explanations of intelligent behavior.

There has been a strong interaction among Artificial Intelligence, Cognitive Psychology and Linguistics (Barr and Feigenbaum, 1981), where people have developed computational (algorithmic) models of processes that they thought could lead to understanding a given string of symbols, or to express a given idea with a proper sentence. In this case, the underlying causal relationship that has been considered is that mental processes generate algorithms that are translated into linguistic expressions, which in turn can be the cause of the display of a specific behavioral response.

In order to link Neuroscience with other cognitive sciences (e.g., Cognitive Psychology, Robotics, and Artificial Intelligence), there has been a tendency to interpret some neuroscientific results as if they could be understood as the outcome of the execution of a logical sequence of steps (Iberall and Lyons, 1984). That is, it is possible to give an algorithmic-like description of the relationship between the parametric conditions needed to elicit a specific behavior and the behavior itself, as if the corresponding algorithm were the "mental" cause of the behavior.

Empirical studies of CNS dynamics have given us conclusive proof that sensory signs (e.g., visual, tactile and auditory) can in fact be translated into regular patterns of neuronal activity (behavior) (Wuriz; Romo; Martinez-Urieegas; Shamma; this volume). Neuronal networks seem to be "programmed" to translate or to process the information of the outside world in a very efficient way. Nevertheless, it is obvious that the outside world is not the only source of signs that have to be processed by the CNS. In fact, the failure of behaviorism was due to the assumption that the same sequences of outside signs (same stimuli) had always to be translated by the system to the same sequence of behavioral activity (same response). Sensory stimuli, in this context, were thought to be the only source of information that had to be processed by the system. We now know that in order to perform a behavioral activity, the CNS seems to require the integration of internal as well as external signs (see Houk, this volume). This means that even if it is possible for neurons to be like very small machines that are capable of translating different signs into

specific activities, the integration of different activities into logical sequences is not fixed in advance. Signs and their translations into behavioral activities need to be put together by the system in accordance with certain rules, which may seem that the CNS is able to build algorithms. That is, it seems that the CNS can build formal devices that provide the correct instructions to perform behavioral activity. This may sound correct, but some methodological inconsistencies are at work here.

Although it is always tempting to mix in teleological descriptions while using purely mechanical explanations, if the aim of scientists is to explain how the CNS works in terms of its mechanisms, it is incoherent to sustain that the CNS "plans" its actions in advance (Arbib; Georgopoulos; Rudomin; this volume). This is only a very confusing way to express our ignorance with respect to how the CNS has to work in order to integrate coherent behavioral activities mechanically. Nevertheless, such phrases or conceptual (theoretical) constructions have the enchantment to make us believe that we are giving answers to those very points that need to be explained; in this case, the relationship between artificial (computational) and natural (neuronal) models of intelligent behavior.

In order to establish whether or not we can develop a global unifying methodology for Cognitive Science, in particular for neuroscience and computer science (Dufee; Arbib; this volume) we must first notice that the disciplines use the information processing paradigm to explain behavior at different levels of analysis. The information processing paradigm can be understood as an approach that follows a functional analysis of behavior (Malcolm, 1964). This kind of approach holds that all the elements that comprise a behavioral pattern can be analyzed in terms of functions that express causal relationships between specific parametric conditions and the activation of specific behavioral responses (Arbib, this volume; Cervantes-Pérez, 1989). For example, if we place a praying mantis on a platform with a certain height in front of a higher pole, at a distance longer than three times the length of the animal's body (parametric experimental conditions), then the praying mantis will execute a ballistic jump from the platform to the pole (corresponding behavioral response). Once the behavioral reference has been established, we may search for the subserving mechanisms. It is important to recognize that under the information processing paradigm, one task is to study the relationship between stimulation conditions and the corresponding behaviors, and another task is to identify the underlying mechanisms that have to be active to produce those specific lines of behavior. The synthesis of artificial systems that display some kind of intelligent behavior relates to the first task; whereas

neurophysiological studies are more related to the second. Obviously, both tasks can be closely interrelated, as long as we are able to recognize their different realms of action, and their different perspectives to explain the same behavior.

In the functional analysis, the control of the elements that comprise a behavioral pattern implies the description of the different functions that express causal relationships between specific parametric conditions and the activation of specific behavioral responses. We can describe this relationship in terms of the transmission and integration of coded signs among different processing elements. In order to solve the problems that may arise in this kind of explanation, we have available three kinds of formal techniques: Information Theory, Control Theory and the Theory of Computation.

Additionally, a functional analysis does not identify the explanation of behavior with specific mechanisms, rather it gives us the advantage of associating the same functions with different kinds of mechanisms. Therefore, we may start with the description of a specific behavior and then work towards the identification of the underlying mechanisms, whether they are neuronal mechanisms in the study of living animals, or electrical and mechanical devices in the construction of complex automata. That is why Brain Theory has been so successful. On the one hand, a theory-experiment cycle can be established to get a better understanding of a specific behavioral phenomenon, by using experimental data to develop theoretical models conformed by dynamic functions. These functions can be translated into an algorithm to be programmed in a computer. Computer simulations are then used, first, to test hypotheses generated experimentally by identifying the parametric conditions under which the dynamics of the biological phenomenon under study is reproduced, and, second, to predict new hypotheses by running computer simulations in which these conditions are changed (see Arbib; Cervantes-Pérez; Marder; Martínez-Urriegas; this volume). On the other hand, we may abstract functional principles from the analysis of biological systems that may be useful, once they are properly formalized, in the construction of automata that display complex behaviors (Arbib, 1989; Houk; this volume).

Acknowledgments. This paper was partially supported by DGAPA under Grant IN-6007 89. We like to thank Dr. Alejandro Tomasini for his help and valuable comments on the manuscript.

References

- Arbib, M.A., 1989, *The Metaphorical Brain 2: Neural Networks and Beyond*. Wiley-Interscience, New York.
- Jarr, A., and Feigenbaum, E., 1981, *Handbook of Artificial Intelligence*, Volume I. Heuristech Press, Stanford, California.
- Block, N. and Fodor, J., 1980, What Psychological States Are Not, *Readings in Philosophy of Psychology*, Vol. 1, (Ned Block, Ed.), Harvard Univ. Press.
- Churchland, P.S., and Churchland, P., 1990, Could a machine think?, *Artificial Intelligence: a debate*, *Scientific American*, January.
- Cervantes-Pérez, F., 1989, Schema theory as a common language to study sensorimotor coordination. In: *Visuomotor Coordination: amphibians, comparisons, models and robots*, (J.-P. Ewert and M.A. Arbib, Eds.), Plenum press, New York, pp. 421-450.
- Cohen, P. and Feigenbaum, E., 1982, *The Handbook of Artificial Intelligence*, Volume III, Heuristech Press, Stanford California.
- Fischler, M., and Firschein, O., 1987, *Intelligence: the eye, the brain, and the computer*. Addison Wesley, Massachusetts.
- Fodor, J., 1975, *The Language of Thought*, Crowell, New York.
- Fodor, J., 1980, Propositional Attitudes, *Readings in Philosophy of Psychology*, Vol. 1, (Ned Block, Ed.), Harvard Univ. Press.
- Georgopoulos, A. P., Lurito, J. T., Penick, M., Schwartz, A. B., and Massey, J. T., 1989, Mental Rotation of the Neuronal Population Vector, *Science*, 243:234-236.
- Forall, T., and Lyons, D., 1984, Towards perceptual robotics. COINS Tech. Rep. 84-17, University of Massachusetts at Amherst, Mass.
- Malcolm, N., 1964, Behaviorism as a Philosophy of Psychology, in *Behaviorism and Phenomenology*, (T.W. Wann, Ed.), Phoenix Books, The University of Chicago Press.
- Putnam, H., 1960, Minds and Machines, in *Dimensions of Mind: A Symposium*, (Sidney Hook, Ed.), New York Univ. Press.
- Fylyshyn, Z., 1973, What the mind's eye tells the mind's brain: A critique of mental imagery, *Psychological Bulletin*, 68:1-24.
- Fylyshyn, Z., 1980, Imagery and Artificial Intelligence, *Readings in Philosophy of Psychology*, Vol. 1, (Ned Block, Ed.), Harvard Univ. Press.
- Fylyshyn, Z., 1990, Computing in Cognitive Science. In: *Foundation of Cognitive Science*, (Michael A. Posner, Ed.), MIT Press.
- Rogers, C.R., 1964, Toward a Science of the Person, in *Behaviorism and Phenomenology*, (T.W. Wann, Ed.), Phoenix Books, The Univ. of Chicago Press.
- Romo, R., and Romo, A., 1950, La Reconstrucción de la Realidad, *Ciencias Cognitivas*, 12 (163):29-36.
- Rosenbluth, A., Wiener, N., and Bigelow, J., 1943, Behavior, purpose and teleology. *Philos. Sci.*, 10: 18-24.
- Searle, J.R., 1990, Is the brain's mind a computer program?, *Artificial Intelligence: a debate*, *Scientific American*, January.
- Simon, H., 1980, Cognitive Science: the newest science of the artificial. *Cognitive Science*, 4:33-46.
- Simon, H., and Kaplan, C., 1990, *Foundations of Cognitive Science. In: Foundations of Cognitive Science*, (Michael A. Posner, Ed.), MIT Press.
- Stallings, N., Feinstein, N., Garfield, J., Emsland, E., Rumbous, D., Weisler, S., an Baker-Ward, L., 1989, *Cognitive Science: An Introduction*. The MIT Press.

METHODOLOGICAL CONSIDERATIONS IN SYSTEM AND COGNITIVE SCIENCES

NYDIA LARA

*Laboratorio de Neurocomputación
Centro de Instrumentos
Universidad Nacional Autónoma de México
A.P. 70-186, México D.F., 04510, México
Fax 550 03 57*

ABSTRACT

The aim of this paper is to explore some of the conceptual and methodological problems that are creating an internal division among those researchers that are exploring the possibility to develop common grounds for the study of intelligent behavior in natural and artificial systems. The point I want to stress is that behind this division lays the idea that intelligent behavior is the effect or external manifestation of an intelligent process. This idea is provoking all kinds of conceptual mistakes, most of all, because under this interpretation it is easy to believe that the phenomenon to be studied it is the intelligence, and not the behavior of the system. The problem with this kind of interpretation is twofold, because, on the one hand, the term 'intelligence' clearly has mental connotation, and like most mental terms, it is not clear what it is and where it could be located, and, on the other had, this misleading term has been introduced in many scientific fields without a technical definition. Both things together are taking us to a situation where people that have been able to control, predict, or reproduce what has been call an intelligent task, either through mechanical devices or through what is known as 'algorithmic' descriptions, tend to believe that they are explaining two things: the causal source of the behavior and the causal source of intelligent processes.

KEYWORD

Behavior, intelligence, mental processes, naturalistic approach, computational approach, System Science, Cognitive Science.

INTRODUCTION

During the last fifty years the development of computational science and engineering is giving rise to the attempt to assemble computational systems capable to accomplish some of the intelligent activities that are performed by living systems. Since then, there is the idea that artificial systems can help us to understand some of those intriguing processes performed by living systems when they interact with each other and with a complex and changing environment. This attempt has open the possibility, foreseen by Norbert Weiner and Ludwig von Bertalanffy, to establish a strong interaction between those fields that study living systems behavior with those that are trying to build artificial systems that behave like them. Nevertheless, this interaction has been confronting conceptual, theoretical and methodological obstacles, mainly because those researchers that are exploring the possibility to develop common ground for the study of intelligent behavior in animals and machines, tend to believe that their subject matter is the study of the intelligence of intelligent systems (Simon and

Kaplan, 1990; Fischler and Firschein, 1987). This apparently innocuous play of word is provoking all kinds of conceptual mistakes, most of all, because under this interpretation underlie the idea that intelligence, and not the behavior of the system, is the phenomenon to be studied. This, in turn, is giving rise to the idea that the behavior exhibited by a system has to be understood as the effect or external manifestation of its intelligence. However, the term 'intelligence' has been introduced in many scientific fields without a technical definition (Lara and Cervantes-Pérez, 1993), and thus, people do not have a clear idea about what is "that" which have to be explained: the causal source of the behavior or the causal source of the intelligent process.

This ambiguity is taking us to a situation where people that have been able to control, predict, or reproduce what has been call an intelligent task, either through mechanical devices or through what is known as 'algorithmic' descriptions, tend to believe that they are explaining two things: the causal source of the behavior and the causal source of intelligent processes. However, this last issue is giving rise to the unexpected result that researchers that are interested in explaining behavior in terms of the physical processes that take place inside of the system, have metaphysical disagreements with people that is more interested in finding the relations that can be established between different kind of effective procedures of behavioral activities (known as algorithms) and the concrete behavior based on them. The first group believe that not only the behavior, but the intelligent process that guides and controls it, can be explained through the understanding of those mechanisms that underly the behavior, whereas, the later group asserts that the intelligent processes that guide and control the behavior are not physical processes, but mental ones.

The point I want to stress in this paper is that when we ascribe purposes, intentions, desires, etc., to the intelligent performance, it is always tempting to try to find an extra process behind the behavior: the one that perform the intelligent act. Unfortunately, to look for ghostly processes guiding the performance give rise to another question: the nature of the thing that is capable to think, plan, decides, etc. Here minds, brains, neurons, computers, and many other mechanisms as well, could be conceived as possible candidates of the causal source of intelligent processes. However, this is not a scientific question, but a metaphysical one. Let us try to find out how this metaphysical question arises.

THE SOURCE OF INTELLIGENCE

Must people believe that purposive behavior is *the mark and criterion of the presence of mentality in a phenomenon* (James, *Principles*, 1890. p.21). Also, must people believe that it is possible to *infer* the "presence" of "mentality" through the observation of the behavior exhibited by the system that is able to perform an intelligent task. This inference can be seen as the strategy that allows us to start from the observable behavior and then try to go behind the scene (Flanagan, 1984, p.183). In the context of this strategy A. Turing, for example, sustains that if the presence of intelligence could be inferred from an intelligent performance, then it was possible to ascribe intelligence to machines that were able to perform intelligent tasks (Turing, 1950). Since then, two things have happened: on the one hand, the construction, either actually or theoretically, of artificial systems with the power of self-replication, self-reproduction, self-regulation, self-repair, etc., is spreading the tendency to try to explain living systems behavior as if they were highly complicated machines. But, on the other hand, the idea

that intelligence can be inferred from and intelligent performance seems to be enough to convince us that if a machine is capable to behave like a living system, it is because somehow the machine is able to think. Both things together are giving rise to the idea that through artificial devices it is possible to understand how the "mind" works, and that through the study of the internal wiring of a living system we can explain how mental processes are implanted or "emerging" from the special organization of mechanical devices. However, both ideas are giving rise to different metaphysical considerations about what is the causal source of intelligent processes: the mind or the complex activity of the internal wiring of the machine. This, in turn, is provoking that different researchers use their data and results in order to support a particular metaphysical point of view about the causal source of intelligent process. The problem here is that the data and results that can be obtained with different research techniques are inconsistent to prove what they want to assert, in the sense that with the same data and results it is possible to defend the plausibility of any metaphysical position about what is responsible for the production of the ghostly processes they want to explain. Let us take an example to illustrate what is at stake.

SYSTEM AND COGNITIVE SCIENCES

Among those researchers that believe that it is possible to gather and to link the different outcomes of those disciplines that study natural and artificial intelligence we find at least two groups:

- a) one that arises from the interaction established between 'theoretical biology, neuroscience, cybernetics, and the computational branch called 'neurocomputation', and
- b) the other that arises from the strong interactions instituted between psychology, linguistic and the computational branch known as 'artificial intelligence'.

To avoid misunderstandings I am going to call the first group 'System scientists', and the second 'Cognitive scientists'.

System scientist (Rosen, 1985; Kampis, 1991; Klir, 1991) depart from the idea that living as well as artificial systems are organized complex systems. For them, it is clear that to understand how a system is able to perform an intelligent task is a matter to understand how its physical components are arranged and interacting with each other in order for the system to be able to achieve a specific goal. In this fashion, systems scientists sustain that to explain how a natural or artificial system behaves as an intelligent being what is important is to understand its engineering details. Additionally, System scientists assert that biological studies play an important role in the development of system technology, in the sense that biological data provide new solution to old problems, and suggest new kind of artificial system inspired in biological ones. In this fashion, for them it is conceivable, first, to assemble an artificial system in such a manner that it can "replicate" the same complex activity that takes place in the internal wiring of a natural system, and second, that this simulation can serve to explain how a system perceives, transforms, or elaborates an external stimulus in order to be able to achieve a particular goal. However, in their terms goals and purposes could be explained only in terms of the internal physical activity of the system.

Fronting this particular way of looking at things, Cognitive scientists allege that the term 'organization' is not related with a particular physical organization, but that this term has to be understood as a cognitive or computational organization. This group sustains against System

scientists, that research in Artificial Intelligence has given enough proofs to support the idea that thinking plays a causal role in the determination of intelligent behavior, and that thinking is not a physical process, but a cognitive one. In fact, for them Artificial Intelligence confirmed the idea that the psychological aspects of intelligent behavior are not related with its physical aspects. The same kind of 'cognitive processes', that is, the same mental or computational processes, could indistinctly be accomplished by systems that, first, clearly are not assembled in the same fashion, and, second, that do not have to perform the same sort of physical processes to be able to achieve the very same cognitive or intellectual job. In this way, Cognitive scientists held that in order to explain how any systems (natural or artificial) behave itself as an intelligent being, the only thing that matters is the study of the intelligent process that defines the behavior to be performed, regardless of the hardware that executes it (Fodor, 1975, 1980; Pylyshyn, 1980; Simon and Kaplan, 1990; Haugeland, 1981; Flanagan, 1991).

It is obvious that we have two sets of areas that are trying to explain the causal sources of intelligent processes from different approaches. System scientists use what is known as the 'naturalistic' approach, maybe because its strong link with biological studies. Cognitive scientists use what has been called the '*artificial*' or '*computational*' approach, probably because, on the one hand, this approach makes use of the software-Hardware distinction derived from Artificial Intelligence to introduce, in a computational language, the mind-body distinction, and, on the other hand, because with this distinction, this approach wants to reject the idea about the possibility that intelligence, as well as other psychological processes could be explained through the so called 'naturalistic approach'.

In fact, the main concern of the naturalistic approach is to find out how a system is or has to be build in order for it to be able to achieve a specific goal. Whereas the main concern of the computational approach is to understand the so called 'cognitive' or 'intelligent' processes that conform the logical structure of intelligent behavior, independent of the nature of the physical system that exhibits this behavior (Simon and Kaplan, 1990, p.2). Through time, this difference has been expressed as a contest between those that study the so called 'intelligent processes' as an outcome of the work of a physical device, and those that study those very same intelligent processes but as an outcome of mental states (e.g., beliefs, desires, goals, hopes, thoughts, intentions, expectations, etc.).

CONCLUSION

The point I want to stress to finish this paper is that the impossibility to decide not only what intelligence is, but where it can be located, respond to the fact that intelligence *per se* is not a phenomenon that can be explained by itself. In fact, it seems that the only way that we are able to maintain that the behavior of any system is "intelligent", is through the performance of the system. This means that 'intelligence' refers more to our own way to evaluate the outcome of the performance, than to an inferred mental process. However, it is clear that when we are willing to ascribe purposes, intentions, desires, etc., to the performance, it is always tempting to try to find an extra process behind the behavior: the one that performs the intelligent act. This introduces us into metaphysical considerations that deviate the scientific research in favor of those sorts of questions that are more related with the philosophy of mind than with science itself.

BIBLIOGRAPHY

- Fleisher, M. and Firschoin, O. (1987); *Intelligence: The Eye, the Brain, and the Computer*, Addison-Wesley Pub. Co.
- Flanagan, O. (1984); *The Science of Mind*, 2th. Ed., MIT Press, London, England.
- Fodor, J. (1975); *The language of thought*; Crowell, New York.
- Haugeland, J. (1981); "Mind Design"; *The Nature and Plausibility of Cognitivism*; MIT Press, Cambridge, Mass.
- James, W. (1890); *The Principles of Psychology*, 3 Vols., Harvard Univ. Press, Cambridge, 1976.
- Kampis, G. (1991); *Self-Modifying Systems in Biology and Cognitive Science*; IFSR International Series on System Science and Engineering, Vol. 6; Pergamon Press, Great Britain.
- Klir, G. (1991); *Facets of System Science*; IFSR International Series on System Science and Engineering, Vol. 7; Plenum Press, New York-London.
- Lara, N. and Cervantes-Pérez, P. (1993); "Methodological Considerations in Cognitive Science"; *Neuroscience: From Neural Networks to Artificial Intelligence*, Vol. 4, Research Notes in Neural Computing; Springer-Verlag, USA.
- Lyubskaya, Z. (1980); *Computation and Cognition*; MIT Press, Cambridge, Mass.
- Loose, R. (1985); *Anticipatory Systems*; IFSR International Series on System Science and Engineering, Vol. 1, Pergamon Press, Great Britain.
- Simon, H.A. and Kaplan, C.A. (1990); "Foundations of Cognitive Science"; *Foundations of Cognitive Science*, ed. by Michael I Posner, MIT Press, London, England.
- Furing, A.M: (1950); "Computer machines and intelligence", "Mind", Vol. LIX:236

CONDUCTISMO Y PSICOLOGIA COGNITIVA

NYDIA LARA ZAVALA

FRANCISCO CERVANTES-PEREZ

INTRODUCCION

Aunque nos gusta pensar en el conocimiento humano como el resultado de un proceso acumulativo y lineal, hay ocasiones en las que resulta muy difícil reconocer el valor de lo que de hecho es un gran avance intelectual y que permite abordar un viejo problema con una perspectiva nueva. Más aún, hay inclusive casos en donde el intelecto retrocede ante sus propios logros y propone como novedoso el tratamiento de un problema previamente superado. El surgimiento de la psicología cognitiva, por ejemplo, es uno de esos casos en donde investigadores formados en la línea del conductismo optan por rescatar la idea de que la psicología no podría ser concebida como una ciencia completa mientras ésta no diera cuenta de los denominados "procesos internos" de la conducta. Según este grupo, distintos estudios conductuales habrían dado pruebas contundentes de que, si bien es verdad que ciertas condiciones de estimulación provocan en el organismo ciertos patrones conductuales regulares, es obvio que el mundo externo no es la única fuente de información que procesa el organismo. Hay casos en donde las mismas condiciones paramétricas de estimulación externa no producen el mismo patrón conductual: importantes factores internos, o lo que los etólogos modernos actualmente denominan "estados motivacionales", parecen jugar un papel decisivo en la manera como se desencadena una determinada conducta. Por lo que la conducta no es, como pensaban los conductistas, sólo el resultado de la estimulación externa: su estudio y explicación requerirán más bien de la consideración de los factores internos que afectan al organismo. O sea, el organismo parece estar "programado" para responder de cierta forma a los estímulos externos, sólo que factores internos juegan un papel modulador (y mediador) en la emisión de la conducta.

Vale la pena notar que, justo al mismo tiempo que surgía la psicología cognitiva, (cargada de pretensiones concernientes a la fundamentación de una supuesta "ciencia de la mente"), Skinner advirtió que el retorno a un mentalismo, aparentemente apoyado en los avances de las neurociencias y el desarrollo de lo que hoy conocemos como ciencias de la computación, hacía correr un doble riesgo. No sólo podía suceder que se redujera la psicología a una rama más de las neurociencias, sino que también se corría el peligro de introducir en las neurociencias mismas y en las nascentes ciencias de la computación problemas propios de la filosofía de la mente y que no compete a una ciencia experimental, como lo es la psicología, trabajar. El objetivo de este ensayo consiste, pues, en mostrar que, tal como Skinner lo sospechó, el mentalismo de la psicología cognitiva, o la línea de investigación que se auto-denomina "psicología del procesamiento de información" o "ciencias cognitivas", no sólo no le ha aportado nada nuevo a la psicología, sino que la ha

empobrecido, al mismo tiempo que ha introducido tanto en las neurociencias, como en las ciencias de la computación, la lingüística, y, en general, en todas aquellas disciplinas que de uno u otro modo se ocupan del fenómeno de la conducta "inteligente", la ilusión de que la ciencia puede genuinamente estudiar los procesos "mentales" que acompañan a esta clase de conducta. En este trabajo me propongo simplemente llegar al punto desde donde se pueda apreciar que lo que sostienen los cognitivistas como una posición novedosa frente al conductismo es una postura que hace cuarenta años fue catalogada por Skinner como un error típico y en el que una y otra vez caen aquellos investigadores que piensan que es posible estudiar las denominadas "causas internas" para dar cuenta de la conducta inteligente.

I

LA CIENCIA DE LA CONDUCTA

La experiencia parece mostrar que tratar de estudiar las causas de la conducta buscándolas en el interior del organismo es una práctica que más que ayudar a comprender la conducta tiende a oscurecer su explicación. Es claro que una conducta tiene causas internas (fisiológicas), pero lo que realmente parece importante es ante todo discernir qué tanto contribuye la investigación de las causas internas a la consolidación de la ciencia de la psicología.

Contrariamente a lo que sostienen muchos de los críticos del conductismo, el programa que éste representa de ninguna manera se puede interpretar como una propuesta que niegue la existencia de causas internas (llámeseles neurofisiológicas o mentales). Lo que el conductista propone es mucho más interesante: se trata de deslindar el objeto de estudio de la psicología de aquellas disciplinas que buscan la explicación de la conducta en el interior del organismo. Por ello, más que ofrecer una teoría de la psicología, lo que el conductista ofrece es una metodología para abordar temas de psicología tratando de esclarecer cuál es el objeto real de estudio de dicha ciencia y cuál es el mejor camino para aproximarse a él. Y parece innegable que la psicología tiene que empezar por reconocer que lo único que puede estudiar es la conducta del organismo, lo cual no requiere del conocimiento ni de los mecanismos ni de los estados internos o de lo que confusamente se cataloga como estados "mentales" que la acompañan.

La comprensión del programa conductista hace factible elevar objeciones en contra de programas cognitivistas. Por ejemplo, se puede argumentar que, dado que "lo mental" por definición carece de dimensiones físicas, entonces se vuelve con suma facilidad o bien a la postulación de la actividad de una entidad fantasmagórica imposible de conocer científicamente o bien la actividad del sistema nervioso. En ambos casos hay problemas insolubles ya que, por un lado, la mente puede aparecer prácticamente en cualquier lado y sustituir con pseudo-explicaciones los temas que se quieren conocer. Por otra parte, aunque en cierto sentido la neurofisiología puede dar cuenta de los mecanismos o procesos fisiológicos subyacentes a la conducta, sus resultados, aunque complementarios a los resultados de las investigaciones de la ciencia de la conducta, no eliminan la independencia

de las respectivas disciplinas ni las hace una reducible a la otra. Debe quedar claramente establecido que una cosa es estudiar la relación entre las condiciones de estimulación y la conducta correspondiente que dicha estimulación desencadena y otra es tratar de identificar los mecanismos neuronales que se activan cuando se desencadena cierto tipo de conducta. El estudio de la conducta puede generar investigaciones concernientes a los mecanismos del sistema nervioso que le subyacen, pero sólo por un error conceptual se puede llegar a suponer que el conocimiento de la actividad de los mecanismos neuronales equivale al conocimiento de lo que ocasiona la expresión significativa de determinada conducta. El neurofisiólogo puede llegar a decirnos con exactitud qué mecanismos se activan cuando un determinado movimiento físico se produce, pero esto no significa que el mecanismo que se activa sea la causa conductual, esto es, la causa relevante de esa conducta que tiene como manifestación a ese movimiento físico, independientemente de que sea posible provocar ese particular movimiento totalmente fuera de contexto. El psicólogo estudia el movimiento físico que exhibe el organismo en contexto, y es el contexto lo que le da a esa conducta su particularidad psicológica. Lo que tiene que considerar el estudio de la conducta contextualizada es cómo responde el organismo a los estímulos del medio. Ahora bien, esto no tiene nada que ver con el estudio de los mecanismos que se activan para que el organismo pueda moverse de la manera como se mueve. Esta es la diferencia entre la psicología y las neurociencias: la primera busca en el exterior los estímulos y las reacciones del organismo ante ellos e interpreta lo que resulta de dicha relación lo que hace posible la conducta; el segundo busca en el interior los mecanismos que la permiten los movimientos físicos del organismo ante determinado estímulo. Es la capacidad que tiene el neurofisiólogo de generar los movimientos físicos lo que hace que muchos estudiosos de la conducta confundan los aportes de la neurofisiología con resultados de psicología. Pero lo que deseo sostener es que lo que hace el neurofisiólogo no es dar una explicación de la conducta, sino sólo de los mecanismos que la permiten. Por otra parte, habría que observar que lo que les permite a los neurofisiólogos explicar lo que internamente le acontece al organismo cuando se activa una conducta es la conducta que exhibe el organismo cuando ciertas condiciones externas lo afectan. Expliquemos esto con un breve ejemplo.

En el estudio de la conducta normalmente se ve a esta última como el resultado de tres factores causales, a saber: a) un factor externo que opera en el organismo, como puede ser la privación alimenticia; b) una condición interna, como puede ser el estado fisiológico o psíquico de hambre; y c) la conducta, que en este caso sería la acción de comer. Sin duda el conocimiento de las condiciones internas puede llevarnos a predecir la conducta del organismo; pero lo que hay que señalar es que exactamente la misma condición interna normalmente sólo se presenta cuando un factor externo opera en el organismo. El fisiólogo puede decirnos que un animal come cuando los niveles de nutrientes bajan de cierto valor umbral, pero la explicación de la causa interna requiere de la consideración del factor externo, esto es, del conocimiento de que el animal come o tiende a comer cuando los niveles de nutrientes están bajos. Sin embargo, cabe señalar que no es absurdo pensar que un animal puede comer sin que sus nutrientes estén bajos o que un animal no coma aunque haya una disminución de sus niveles de nutrientes. Esto indica simplemente que la conducta del organismo no puede explicarse sólo sobre la base de lo que le acontece en su interior. De alguna manera, lo que se tiene que reconocer es que más tarde o más temprano tenemos que

llegar al punto donde se tiene que tomar en cuenta que el organismo se comporta como se comporta porque hay factores externos que operan en él (e.g., históricos, sociales, ambientales, etc.). El fisiólogo estudia las estructuras y los procesos mecánicos sin los cuales el organismo no podría comportarse como se comporta, pero al psicólogo le toca dar cuenta del comportamiento del organismo analizando la manera como éste reacciona conductualmente cuando queda expuesto a los embates de los factores externos que lo afectan y que lo conducen a actuar de cierta manera ante ciertas circunstancias. La baja de nutrientes puede impulsar a un animal a comer, pero el estudio de la conducta de comer tiene que considerar muchas más cosas que una simple baja de nutrientes.

II LA CIENCIA DE LA MENTE

Diversos experimentos muestran que un organismo puede responder a una situación ambiental específica con toda una variedad de líneas de conducta. Curiosamente, este hecho es el que ha utilizado la psicología cognitiva para postular que la conducta es el resultado de lo que ellos consideran un ciclo de percepción-acción. Dicho ciclo se explica bajo la guía de lo que actualmente se denomina "paradigma de procesamiento de información", el cual ha sido interpretado por los psicólogos cognitivistas como un proceso que requiere la intervención de cuando menos cinco fases: a) la estimulación externa; b) la traducción de los signos externos en señales nerviosas; c) la interpretación que el organismo hace de esas señales nerviosas en símbolos para conformar lo que los psicólogos cognitivistas denominan un "modelo interno del mundo externo"; d) la planeación de la conducta a seguir considerando tanto lo que sucede en el mundo externo como lo que sucede en el interno; y e) la acción o conducta concreta que realiza el organismo conforme a un plan o lo que en términos computacionales se denomina un "algoritmo".

Como la psicología cognitiva basa sus explicaciones del ciclo percepción-acción tomando como punto de partida las estructuras de procesamiento de información (i.e., estructuras que permiten la manipulación de símbolos), de manera natural desde el principio se estableció una estrecha relación entre los psicólogos cognitivistas y quienes trabajaban en las áreas dedicadas a la investigación de la denominada "inteligencia artificial" - la rama de las ciencias de la computación dedicada al estudio de técnicas para construir sistemas computacionales (i.e., autómatas en el sentido de la máquina de Turing). Un autómata es una máquina capaz de exhibir algunos aspectos de conductas globales tales que, cuando son llevadas a cabo por seres vivos permiten inferir que exigen para su realización de eso que llamamos "inteligencia".

Como normalmente se piensa que la inteligencia es un proceso mental y que ésta juega un papel activo en la producción de la conducta es fácil suponer que la conducta inteligente es el producto de lo que la inteligencia determina.

En la construcción de autómatas que lleven a cabo conductas inteligentes podemos detectar dos líneas de investigación bien definidas: una que trata de transformar las explicaciones mecánicas que se obtienen de los estudios de la dinámica del sistema nervioso

central (SNC) en modelos artificiales o teóricos que simulen la ejecución de la conducta inteligente (Arbib, 1992, Barto, 1992, Zak, 1992); y otra que trabaja con la herramienta computacional para construir sistemas (algoritmos) que permitan en principio llevar a cabo conductas inteligentes, independientemente de cómo las realicen los seres vivos (Lee, 1992, Marroquin, 1992, Riseman, 1992). En las dos líneas de investigación mencionadas hay casos en donde encontramos que la atención se centra básicamente en la simulación de los movimientos corpóreos asociados con una acción inteligente, pero hay también casos en donde los investigadores intentan modelar, además de los movimientos corpóreos, procesos denominados "racionales", como son los juegos, la solución de problemas, el reconocimiento de patrones, etc.

Dependiendo de las diferentes técnicas de investigación que se utilicen en la construcción de un autómata, es posible encontrar básicamente cuatro tipos de modelos que explican o simulan los llamados procesos "internos" de lo que se concibe como la conducta "inteligente". Estos son: a) los biológicos (neuronales); b) los formales (lógico-matemáticos); c) los computacionales (conceptuales); y d) los artificiales (máquinas). Diré unas cuantas palabras al respecto.

Los modelos biológicos normalmente se constituyen sobre la base de los mecanismos neurofisiológicos que se conocen cuando el animal (incluyendo al humano) realiza alguna tarea "inteligente" (i.e., captura de presas, reconocimiento de patrones, discriminación de estímulos, etc.). Los otros tres analizan la conducta inteligente no como un proceso físico, sino como un procedimiento efectivo que puede ser modelado haciendo abstracción del "cuerpo" o de los mecanismos que subyacen a su ejecución. Esto es lo que ha llevado a interpretar la construcción de algoritmos como el estudio científico de los procesos mentales. En efecto, el algoritmo parece contener la realización de dos tipos de procesos, uno relacionado con la lógica de la secuencia de movimientos que tienen que ejecutarse para llevar a cabo una tarea inteligente y otro relacionado con la posible traducción de esa secuencia lógica (racional) en una serie de movimientos físicos concretos.

Podría replicarse, sin embargo, que los dos pasos que contempla una descripción algorítmica (la secuencia lógica de la conducta y su traducción en movimientos físicos) en realidad no son sino una única descripción, una re-descripción de uno y el mismo fenómeno (la descripción de la ejecución de la conducta y la observación de la ejecución de la conducta), por lo que sólo por error se puede suponer que la descripción de un procedimiento efectivo pueda ser la causa del procedimiento efectivo que se describe. No obstante, la formulación de la explicación de un proceso en términos de dos descripciones nos lleva a suponer que una de las descripciones es la verdadera explicación y que la otra es redundante. Así, fácilmente se llega a la conclusión de que dado que conocemos la causa interna de una conducta, entendida como un procedimiento efectivo, ya no hay nada más que buscar.

Los algoritmos describen procesos que pueden ser ejecutados, pero ellos no son ni acciones ni cosas (mentales o físicas) que puedan por ellas mismas generar acciones. Lo que podría decirse al cognitivista es que él está interpretando al algoritmo como si fuera una

cosa que guía una conducta, por lo que es fácil suponer que la conducta puede ser causada por un algoritmo. Sin embargo, si se comprendiera que los algoritmos no son cosas que guían conductas o que las planean, sino que son simples sistemas de reglas para describir conductas que pueden ser ejecutadas por cosas (animales, máquinas o humanos), entonces sería más fácil comprender que el algoritmo no acarrea consigo ningún elemento de carácter mentalista.

Ahora bien, pese a sus confusiones, el papel de la psicología cognitiva es importante, puesto que actualmente parece central en los intentos por relacionar áreas aparentemente tan dispares como son las neurociencias, la ingeniería de control, la inteligencia artificial, la robótica, la lingüística, etc. Su contribución en la conformación de lo que ahora se denominan las "ciencias cognitivas" es, empero, más bien pernicioso. La idea es la siguiente: si bien es cierto que la psicología cognitiva ha logrado relacionar a las distintas disciplinas que se supone que estudian las causas de la conducta denominada "inteligente", también lo es el que lo ha logrado introduciendo una enorme gama de postulados metafísicos. En particular, ha re-introducido, de manera totalmente gratuita, supuestos y principios de corte mentalista en los estudios de la dinámica de la actividad neuronal y en los programas computacionales.

De hecho, como resultado de la labor de los cognitivistas, es ahora muy común encontrarse a neurocientíficos y computólogos describiendo la actividad de una neurona, de un cerebro o de una máquina recurriendo a un lenguaje mentalista, lo cual no sólo no le agrega nada a la explicación, sino que introduce en objetos extraños a la psicología la gama de problemas que el psicólogo o el filósofo de la mente estaban tratando de eliminar o esclarecer. El corolario de esto es que gran parte de la literatura concerniente a las denominadas 'ciencias cognitivas' gira alrededor de cuestiones que competen más a los estudios relacionados con la filosofía de la mente que a la ciencia misma de la psicología. La importación de problemas metafísicos se logró, además, complicando enormemente los temas, puesto que sería función de la psicología cognitiva reunir los data de diversas ciencias (la teoría de la computación, la lingüística, etc.) en un todo homogéneo. Esto ha propiciado recurrir a pseudo-explicaciones que sólo en apariencia conjugan el interés de disciplinas que no estudian lo mismo. Por ejemplo, es gracias a su influencia que ahora resulta de lo más normal hablar de la percepción, memoria, imaginación, recuerdos, etc., no sólo de los humanos, sino también de los cerebros, de sus neuronas, de las computadoras, etc., etc. Esto ha dado pie a que se presenten como genuinas cuestiones de la ciencia intratables y tradicionales problemas de corte más bien filosófico que científico, como el problema mente-cuerpo. Se introducen así toda clase de teorías metafísicas para apoyar o rechazar la posibilidad o imposibilidad de cerebros o máquinas pensantes (un ejemplo claro de este tipo de discusión se encuentra en *Artificial Intelligence: A debate*, Searle, and Churchland and Smith-Churchland, 1990).

Ahora bien, los cognitivistas nos ha llevado a creer que, para poder dar cuenta de una conducta inteligente, la razón juega un papel importante en su producción. Esta forma de entender el término "razón", sin embargo, siempre acaba por hacernos pensar que estamos hablando de un proceso mental como la causa de toda clase de conducta inteligente. Sin embargo, cabe señalar que los diferentes modelos de la conducta 'inteligente' utilizan el

término "razón" para referirse solamente al conocimiento del resultado que se obtiene o se puede obtener de la ejecución de una actividad física específica. Por lo que el significado del término "razón" en esta clase de modelos no puede ser entendido como un término que apunte hacia algún tipo o clase de proceso mental. Las causas del comportamiento para aquellos que estudian o construyen modelos 'cognitivos' se basan en los análisis conductuales y estos no tienen nada que ver con entidades fantasmagóricas ni con los mecanismos que la subyacen.

Por otro lado, la liga entre psicología cognitiva e inteligencia artificial dio lugar a que se creyera que la computadora podía ser concebida como un buen modelo de la mente, sugiriendo que la mente, al igual que la computadora, debe procesar información bajo el paradigma serial, esto es, debe ejecutar una serie de operaciones, pero una a la vez, para conseguir el objetivo de la conducta.

Como el procesamiento lógico y su seguimiento o traducción en movimientos físicos son aparentemente dos actividades distintas, los psicólogos cognitivistas han interpretado al algoritmo que constituye a este tipo de modelos como el modelo de la clase de relación que se supone debe darse entre las actividades denominadas "mentales" y las acciones físicas concretas que se ejecutan conforme a un plan previamente concebido.

Lo que sostiene el cognitivista es que la construcción de la secuencia lógica (racional) que determina la conducta a seguir es distinta e independiente de la ejecución física de esa secuencia lógica, por lo que se puede sostener que una y otra evidentemente no pertenecen al mismo reino. La primera es una actividad conceptual (mental), mientras que la segunda es, sin duda, una acción física que sigue paso a paso las condiciones que le impone la primera. Esto es lo que ha dado lugar a que se postule que los procesos mentales son como programas computacionales que una vez que son concebidos por la mente pueden ser correctamente ejecutados por el cuerpo. Como se puede apreciar, este tipo de interpretación introduce en las ciencias de la computación un dualismo metafísico que asume que el algoritmo no sólo es el modelo del proceso mental, sino la causa del movimiento corpóreo.

Ahora bien, aunque Skinner no trabaja, hasta donde nosotros sabemos, la noción de algoritmo, en sus escritos se puede encontrar una poderosa crítica a la interpretación cognitivista de los procedimientos efectivos que modela la inteligencia artificial. Lo que nosotros suponemos que Skinner diría es que lo que confunde al cognitivista es que lo que el algoritmo modela es el proceso a través del cual se puede obtener cierto tipo de resultado cuando se ejecuta una conducta siguiendo cierta secuencia de acciones. Pero para Skinner sería más que evidente que algoritmo no describe un proceso mental, sino la manera como se ejecuta una tarea determinada paso a paso. Esto significa que lo que los cognitivistas interpretan como un proceso mental no es más que la descripción redundante de la conducta que ejecuta el organismo, ya que el algoritmo lo que describe es sólo la manera como se ejecuta una conducta no lo que causa la ejecución de la conducta.

BIBLIOGRAFIA

Arbib, M. A., *Cooperative distributed problem solving between (and within) intelligent agents*, Neuroscience: From Neural networks to Artificial Intelligence, Springer-Verlag, N. Y., 1993.

Barto, A. G. & Gullapalli, V., *Neural networks and adaptive control*, Neuroscience: From Neural networks to Artificial Intelligence, Springer-Verlag, N. Y., 1993.

Lee, M., Anderson, C. H., Weinder, R. J., *State of the art in image processing*, Neuroscience: From Neural networks to Artificial Intelligence, Springer-Verlag, N. Y., 1993.

Marroquín, J. L., *Computational vision: A problematic view of multi-module paradigm*, Neuroscience: From Neural networks to Artificial Intelligence, Springer-Verlag, N. Y., 1993.

Riseman, E. M. & Hanson, A. R., *Where does the knowledge base come from?*, Neuroscience: From Neural networks to Artificial Intelligence, Springer-Verlag, N. Y., 1993.

Searle, and Churchland and Smith-Churchland, *Artificial Intelligence: A debate*, Scientific American, 1990

Skinner, B. F., *Science and Human Behavior*, The Macmillan Company, USA, 1953.

Zak, M., *Dinamical approach to collective brain*, Neuroscience: From Neural networks to Artificial Intelligence, Springer-Verlag, N. Y., 1993.

Rosch, E. y Marvin, C. (1975): 'Family resemblances: studies in the internal structure of categories', *Cognitive psychology* 8, 382-439.

Rosch, E., Simpson, C. y Miller, B.S. (1976): 'Structural basis of typicality effects', *Journal of Experimental Psychology: Human Perception and Performance* 2, 491-502.

Roth, E.M., Shoben, E.J. (1983): 'The effect of context on the structure of categories', *Cognitive psychology* 15, 346-378.

Rumelhart, D., Smolensky, P., McClelland, J. y Hinton, G. (1986): 'Schemata and sequential thought processes in PDP models', en *Parallel Distributed Processing*, vol. II, MIT/Basilford, Cambridge, MA, pp. 7-57.

Smart, J.J.C. (1965): 'The Methods of ethics and the methods of science', *Journal of Philosophy* 62, pp. 344-349.

_____ (1990): 'Integrity and squeamishness', en *Utilitarianism and its Critics*, J. Groves (comp), MacMillan, Nueva York, pp. 170-174.

Smith, E.A. (1990): 'Categorization', en *Thinking: An Invitation to Cognitive Science*, vol. III, D. Osherson y E. Smith (comps), MIT, Cambridge, MA, pp. 33-53.

Smith E.A. y Medin, D.L. (1981): *Categories and concepts*, Harvard University Press, Cambridge, MA.

Smolensky, P. (1988): 'On the paper treatment of connectionism', *The Behavioral and Brain Sciences* 11, pp. 1-74

Stich, S.P. (1990): *The Fragmentation of Reason*, Bradford/MIT Press, Cambridge, MA.

_____ (por aparecer): 'What is a Theory of Mental Representation?'

Tversky, A. y Gati, I. (1978): 'Studies of Similarity', en *Cognition and Categorization*, E. Rosch y B. Lloyd (comps.), Lawrence Erlbaum Associates, Hillsdale, N.J.

White, A. R. (1975): 'Conceptual analysis', en *The Owl of Minerva*, C.J. Bontempo y S.J. Odell (comps), McGraw Hill, Nueva York.

Williams, B. (1970): 'The self and the future', *Philosophical Review*, vol. LXXIX (2), 161-180.

Wittgenstein, L. (1953): *Philosophical Investigations*, Mac-Millan, Nueva York.

Nydia Lara Zavala

Francisco Cervantes-Pérez

Es factible sostener que la filosofía puede jugar tres papeles distintos en relación con la gestación y el desarrollo de una ciencia:

- a) Puede enfrentarse a un problema filosófico tradicional para deslindar terrenos y permitir así una formulación clara de la clase de preguntas que pueden ser abordadas por la ciencia,
- b) puede contribuir a la gestación de nuevas herramientas metodológicas para abordar el estudio de una nueva cuestión, o
- c) puede utilizar los recientes avances que se han generado en alguna o algunas áreas de la ciencia para explorar si la nueva evidencia puede ofrecer información que le sea útil para resolver un problema filosófico.

No es exagerado afirmar que durante la gestación y el desarrollo de las ciencias que se ocupan de la conducta (natural y artificial) bajo el paradigma de procesamiento de información, actualmente denominadas 'ciencias cognitivas', la filosofía de hecho ha jugado estos tres papeles en distintos momentos de la historia de dichas ciencias. En este trabajo vamos a tratar de mostrar cómo en efecto la filosofía ha desempeñado estos tres papeles y lo que se puede lograr gracias a ella.

La filosofía ha jugado un papel importante en la conformación de las ciencias cognitivas, en particular gracias al desarrollo y

la aplicación de dos de sus ramas: la lógica y la filosofía de la mente. Desde ambas perspectivas se abordó, sobre todo durante el siglo XIX y, podríamos decir, hasta Frege, lo que parecía ser un mismo problema, a saber, la organización y el contenido del lenguaje, puesto que las así llamadas 'leyes del pensamiento', de las cuales se ocupaban los psicólogos, no eran otra cosa que las leyes lógicas que rigen al lenguaje. El estudio de la lógica y el estudio de la mente parecían, pues, ser dos caras de una misma moneda.

El estudio del lenguaje absorbió la atención no sólo de muchos psicólogos, sino también la de no pocos matemáticos y de algunos lógicos, por la sencilla razón de que, como se sugirió más arriba, el lenguaje era visto como el instrumento que permitía procesar y transmitir un pensamiento objetivo, el cual podía quedar identificado como el resultado de alguna clase de actividad interna al sujeto (mental o física). Sin embargo, el estudio del lenguaje en tanto que expresión de la actividad interna de pensar muy pronto dio lugar al surgimiento de dos nociones antagónicas del pensamiento: la psicologista y la anti-psicologista. En la postura psicologista (iniciada por gente como Fries y Beneke) se concebía al pensamiento como esencialmente una actividad de la mente. La postura anti-psicologista, en cambio, concebía al pensamiento como una especie de cálculo, es decir, como una técnica o procedimiento formal, más que como un producto de la mente. Ahora bien, contrariamente a lo que muchos podrían pensar, este antagonismo no ha perdido vigencia, puesto que la misma tensión entre estas dos nociones opuestas de pensamiento se vuelve a manifestar en la actualidad justamente en las ciencias cognitivas. Esto es lo que ahora pasaremos a hacer ver.

La tradición cartesiana legó a la posteridad la idea de que la esencia de la mente es el pensamiento, por lo que una teoría del pensamiento tenía que estar inevitablemente ligada a una teoría

de la mente. Para los empiristas ingleses, sin embargo, el pensamiento era más bien un poder de la mente, no su esencia, y resultaba de una organización particular de ideas que surgían o provenían de la experiencia, recurriendo para ello a leyes de asociación. De ese modo, para los empiristas una teoría del pensamiento no podía ser equivalente a una teoría de la mente, sino más bien a una teoría de las operaciones "mentales". Por su parte, Kant propuso la existencia de un esquema conceptual *a priori*, gracias al cual se lograba imponer un orden en los datos de la experiencia. De otro modo, lo que normalmente llamamos 'nuestra experiencia' no sería inteligible. En este sentido, para Kant la mente tenía que ser considerada como una estructura regulada por las formas puras de la intuición sensible, las categorías y los *data* generados por su propia actividad. De ahí que, en su caso el conocimiento de las operaciones de la mente sí desembocaba en una teoría de la mente.

Durante el siglo XIX, la psicología era un estudio en el que se mezclaban o amalgamaban diversas nociones de mente. La mayoría de quienes se ocupaban de estos temas parecía aceptar que era el alma lo que dotaba de unidad a los procesos mentales, que eran de lo más variado. No obstante, el material empírico de la psicología tenía inevitablemente que ver con la manera como se asociaban las sensaciones y las ideas en concordancia con ciertas leyes generales comunes a todos los seres pensantes, a todas las mentes. El pensamiento se entendía, por consiguiente, como el resultado de una actividad de la mente lógica y si bien lo que el pensamiento conectaba eran ideas, lo hacía, una vez más, en concordancia con ciertos principios *a priori*. Dichos principios no sólo no eran un producto de la experiencia, sino que eran el presupuesto de la experiencia. La lógica, que también se suponía que se ocupaba del pensamiento, no era considerada como una disciplina completamente autónoma en relación con la psicología. Se creía que el conocimiento de las leyes del pensamiento

era relevante para la justificación de las combinaciones de ideas. Las leyes abstractas del pensamiento eran las mismas para todos, pero las correlaciones de ideas efectuadas con base en dichas leyes variaban de mente en mente. De esa manera, aunque las leyes formales del pensar daban los modos de combinación de ideas, de todos modos no podían decir nada sobre el contenido de pensamientos particulares. Desde esta perspectiva, a la psicología le correspondía dar cuenta no sólo de aquello que se obtenía como medio mental al combinarse las ideas, sino también de la manera como la mente imponía sus leyes de pensamiento para la combinación de las ideas.

Por otra parte, era normal que la forma de abordar el estudio de la organización y los contenidos de la mente se planteara también en términos de introspección. Los psicólogos entrenaban a sus sujetos de estudio a "mirar dentro de ellos" y a reportar verbalmente lo que les pasaba cuando se les pedía que llevaran a cabo actividades de diversa índole. Los reportes verbales se interpretaban como expresiones lingüísticas de los procesos mentales, por lo que el estudio de la mente empezó a concebirse de manera natural en términos de la estructura y las propiedades semánticas del lenguaje. Esto tenía que ser así, puesto que el lenguaje parecía ser como una especie de puente capaz de permitir extraer las leyes abstractas del pensar (dado que el contenido de las actividades mentales a través de símbolos y combinaciones de símbolos representaban ideas) y en parte porque era gracias a los símbolos lingüísticos que los hablantes podían tanto referirse a como describir la conducta y las acciones de las personas. Gracias a la introspección, el sujeto podía dar cuenta no sólo de cómo se realizaban sus actividades mentales, sino también de cómo realizaba o planeaba la ejecución de ciertas actividades conductuales, recurriendo para ello a la generación de modelos internos del mundo externo. La percepción sensible, la fantasía o la imaginación, unidas a los deseos, las creencias,

las intenciones y demás actitudes proposicionales, podían ser analizadas y comprendidas no sólo por medio de las conductas que lleva a cabo un individuo, sino también por medio de la verbalización introspectiva y en la que intervenían todas aquellas variantes que el individuo podía pensar o imaginar sin que necesariamente se actualizaran.

Por la manera como habían sido concebidas las relaciones entre, por una parte, el lenguaje y, por otra, las actividades mentales y conductuales, la mayoría de las investigaciones se abocaron a investigar simultáneamente el funcionamiento y las interacciones de tres clases distintas de procesos:

- a) los mentales, en principio sólo accesible al sujeto cognoscente,
- b) los conductuales, abiertos al conocimiento público, y
- c) los lingüísticos, considerados como el canal a través del cual era posible comunicar y conocer los procesos mentales y vincular así lo público con lo privado.

En este sentido, el lenguaje era considerado como el instrumento público de expresión de las actividades mentales, en tanto que las actividades mentales se concebían como la causa de desempeño tanto de las actividades lingüísticas como de las conductuales.

La convicción de que las actividades mentales no sólo eran la causa sino que eran lo que dirigía la actualización de las otras dos clases de actividades obligó a tratar de dar cuenta, primero del contenido y la organización de la mente para, sobre esa base tratar de explicar la conducta y el lenguaje. En todo caso, lo importante es notar que se tendía a pensar que tanto el estudio del lenguaje como el de la conducta estaban inevitablemente subordinados al de la mente. El problema de este enfoque es que el conocimiento del contenido y de la organización de la mente sólo podía ser abordado o bien a través de la introspección, lo

cual no era un método fiable, puesto que se trataba de un "método" por completo privado, o bien a través del lenguaje. Además, parte de la dificultad era que no se podía determinar con certeza hasta qué punto efectivamente se podía comunicar sin distorsionar, a través del empleo de signos, lo que acontecía en la mente del sujeto. De alguna manera, se pensaba, averiguar esto implicaba poder contrastar las representaciones de una mente con otra para poder determinar si había o no alguna clase de correspondencia entre signos e ideas, lo cual, como ya se dijo, sólo podía hacerse por la vía de la introspección, confinada sólo al conocimiento en primera persona. El acceso a lo mental por la vía del lenguaje ya era en sí mismo un rompecabezas filosófico, pero vale la pena notar que, por la manera como estaban concebidas las conexiones entre los tres tipos de procesos, aun cuando se pudiera dar por sentado que las palabras correspondían de manera sistemática a ideas y que la organización de una oración reflejaba una determinada organización mental, de todos modos quedaba pendiente el otro problema, de explicar cómo se conectaba la actividad mental con la corpórea.

El lenguaje como la expresión de las actividades mentales no sólo exigía la explicación concomitante de tres diferentes clases de procesos distintos (i.e., mentales, lingüísticos y conductuales), sino que también exigía la explicación de sus conexiones. El lenguaje podía ser comprendido como el puente que permitía vincular a las actividades mentales con las conductuales, pero todo parecía indicar que mientras se concibiera el lenguaje como un mero instrumento para la expresión de actividades mentales, el estudio del lenguaje por sí solo no podía proporcionar ninguna clase de conocimiento de lo que acontecía en la mente, además de que tampoco permitía ofrecer una explicación clara de cómo los procesos mentales podían ser traducidos primero en palabras y oraciones y, posteriormente, retraducidos otra vez a ideas o procesos mentales, los cuales (se suponía) son lo que daba origen

a conductas concretas de individuos concretos. Así, este complejo entramado de relaciones e interconexiones entre actividades mentales, lingüísticas y conductuales hace ver que, desde esta perspectiva, se estaba tratando de abordar cuestiones diferentes como si se tratara de una y la misma. En efecto, un asunto era el estudio del lenguaje como una vía para comunicar dos diferentes clases de "actividades", las mentales y las conductuales, y otro era el estudio de las relaciones entre diferentes clases de sustancias, esto es, la mente y el cuerpo. Esta segunda clase de estudio empezaba a revelarse como completamente redundante.

El problema mente-cuerpo es un enigma filosófico tradicional para el cual, como es de esperarse, no hay una solución universal satisfactoria. Hay que notar, por otra parte, que el estudio del lenguaje como estudio del medio idóneo para la expresión de la conducta no equivale al estudio de aquello que la causa. Sea como fuere, el lenguaje aun visto como expresión del pensamiento en el sentido psicologista del término, no puede ser visto como la causa directa de ninguna actividad, sino sólo como la expresión de una línea de conducta posible. Si se dice, por ejemplo, "Trae el lápiz que está sobre la mesa" o "Suma dos y dos", aun cuando es fácil darse cuenta de que la primera oración expresa un movimiento corpóreo, esto es, una actividad física, y la segunda expresa una "actividad" que puede ser catalogada como mental, las dos oraciones aluden a clases de conducta que pueden realizarse sólo a través del empleo de palabras. Ninguna de las dos oraciones apunta a algo más que a una conducta posible, por lo que cualquiera de estas dos oraciones puede ser comprendida simplemente como la expresión de actividades de clases diferentes y que se pueden o no llevar a cabo. En otras palabras, la significatividad del lenguaje se deriva de su poder para expresar diferentes clases de actividades, esto es, conducta específica. Empero, ni lo que expresa el lenguaje ni la clase de

conducta que está indicada en la oración refieren a aquello que los genera. Esto es, cuestiones tales como qué o quién profiere la oración, qué entidad "se esconde" detrás del cuerpo, etc., son cuestiones que dejan de tener relevancia en este estudio, ya que de lo que se trata es simplemente de comprender cómo puede el lenguaje funcionar y dar lugar a actividades conductuales específicas. Desde esta perspectiva, el estudio de la clase de conducta que puede ser expresada por un trozo de lenguaje o el estudio de la manera como una conducta puede desencadenarse a partir de lo que indica una oración, se vuelven objetos de estudio perfectamente objetivos y no involucran ninguna otra clase de consideraciones. Esto es así porque esta clase de estudio sólo versa sobre los efectos que puede generar el empleo de una oración y de las actividades a las que puede dar lugar. Una vez que se aclara que es posible estudiar el lenguaje como si se tratara de una herramienta que contiene mecanismos para expresar alguna clase de actividad o de conducta, el siguiente paso es trazar las líneas de investigación que pueden derivarse de un estudio así. Normalmente, este es el punto donde la ciencia toma la estafeta e inicia su labor por su cuenta.

No estará de más recordar que originalmente se habían planteado simultáneamente alrededor del estudio del lenguaje dos preguntas:

- a) la primera, como ya se dijo, concernía a las leyes del pensamiento, entendidas como leyes de la lógica, y
- b) la segunda se refería a las leyes del pensamiento, entendidas como leyes de la mente.

Ambas clases de investigación convergían en el estudio del lenguaje y, por ende, en el de las relaciones entre el lenguaje y la mente, por una parte, y el lenguaje y la conducta, por la otra. Es claro que difícilmente podía la segunda cuestión separarse

de la primera. Sin embargo, pasaron muchos años antes de que se comprendiera que la psicología era, más que el estudio de la mente, el estudio de la conducta via el estudio del lenguaje y sus descripciones, usos, etc., y que el estudio de las leyes formales del pensamiento desligadas de los simbolismos concretos y sus aplicaciones no pasa de ser un mero juego. El estudio del lenguaje era cada vez más vital y era el canal para la conexión entre el pensamiento, entendido como algo visible o palpable, y la conducta, humana o no.

Sin duda, en los lógicos matemáticos del siglo XIX poco a poco cristalizó la idea de que las leyes del pensamiento, en el sentido anti-psicologista del término, podían concebirse como una serie de combinaciones simbólicas susceptibles de ser conocidas con total independencia de los supuestos contenidos del pensamiento a los que los signos aluden. Sin embargo, los psicologistas tendían a ver en la lógica una disciplina subordinada a la psicología, ya que se pensaba que los signos que se usaban eran forzosamente nombres de ideas y porque se creía que divorciar los signos de sus significados mentales era cancelar el conocimiento de la mente. El problema para ellos era que el recurso a la introspección no podía ser un buen método científico para acceder a nada (los contenidos de la mente o lo que fuere). Pero, además, la única forma de acceso a los contenidos de otras mentes eran las manifestaciones públicas del sujeto y un análisis más minucioso de los reportes verbales que trataban de dar cuenta de las cuasas o motivaciones de sus conductas permitía poner en duda la necesidad de postular que los signos lingüísticos tenían que estar conectados con ideas para ser significativos. En última instancia, cuando se le pedía a un sujeto traer, por ejemplo, un lápiz rojo, lo que menos importaba era si su representación mental de lápiz o de rojo coincidía o no con las representaciones mentales de lápices rojos de otras personas. De hecho, ni siquiera importaba si el sujeto en cuestión efectivamente requería de una

representación mental para llevar a cabo correctamente la tarea que se le indicaba. Lo único que importaba era que el sujeto fuera capaz de procesar la información contenida en la oración para ejecutar correctamente la tarea y realizarla.

Si lo que se ha dicho es acertado, entonces puede apreciarse que el avance consistió en darse cuenta de que la propiedad de las expresiones lingüísticas de expresar una actividad o de desencadenarla poco o nada tiene que ver con lo que acontece en la mente del sujeto, que es independiente de ella. La función del sujeto se puede limitar a simplemente seguir paso a paso y de manera mecánica la regla incorporada en la oración. El contenido de la oración no refiere a sus supuestos significados en la mente del sujeto o al modo como operan las entidades psíquicas, sino sólo a la manera como los signos están conectados unos con otros para expresar o generar una conducta concreta. Esto posibilita dos cosas:

- a) desprenderse de las especulaciones metafísicas, tan apreciadas por los psicólogos, en torno a los contenidos mentales, y
- b) generar una teoría del lenguaje no meramente formal, sino también de sus modos de significación.

La distinción entre contenido mental y contenido lingüístico es una dicotomía que permite orientar la búsqueda de los principios que nos van a permitir eliminar los problemas de las interpretaciones psicologistas y construir las bases metodológicas de toda una serie de trabajos que culmina en la teoría de lo computable y que es, de hecho, la herramienta metodológica que hasta la fecha utilizan las ciencias cognitivas. Esto nos lleva a otro aspecto del tema general abordado en este ensayo.

Durante siglos los filósofos se han enfrentado al problema mente-cuerpo y han tratado de explorar la relación que vale entre la materia y la mente desde prácticamente todas las perspectivas,

sin que haya sido posible obtener una forma de acoplarlas de manera enteramente satisfactoria. De este estudio han brotado las más diversas teorías metafísicas en torno a las clases de entidades involucradas. Por ejemplo, se ha pensado en la mente y en el cuerpo como dos sustancias distintas que pueden tener o no una relación directa, se ha planteado la posibilidad de reducir una a la otra, de convertir a una en causa de la otra, etc. Lo cierto es que hasta la fecha y a pesar de todos los esfuerzos que el desarrollo de cada una de estas posibles combinaciones acarrea, puede afirmarse que la relación mente-cuerpo sigue siendo un problema insoluble. Es, pues, sumamente extraño que el desarrollo de la tecnología computacional y los avances en el área de las neurociencias, actualmente exploradas con entusiasmo por una multitud de filósofos, incite a pensar que las nuevas evidencias obtenidas en estas nuevas líneas de investigación den la pauta para la solución definitiva del problema mente-cuerpo. Nosotros pensamos que esto es un error y que se está sosteniendo en una nueva terminología una tesis que ya había sido superada, viz., la tesis de que el lenguaje no es más que el instrumento para la expresión de algo diferente a él, esto es, algo "mental". Esto, obviamente, tiene que ver con la posición de las ciencias cognitivas en relación con el problema mente-cuerpo. Como se sabe, el nombre 'ciencias cognitivas' denota al grupo de disciplinas que estudian al hombre, a los animales, los cerebros, las máquinas, el lenguaje, etc., tomando como guía el paradigma de procesamiento de información, el cual se deriva directamente de las ciencias de la computación. Las ciencias de la computación surgen del estudio de la relación lenguaje-conducta, pero aunque se reconoce que el lenguaje es capaz de expresar tanto actividades mentales como físicas, no se incluye en esta propuesta ni una solución ni una reformulación del problema mente-cuerpo. Más bien, se promueve una forma de evadirlo. En este caso, lo que se acepta es que el lenguaje puede expresar, a través de palabras

y oraciones, actividades que pueden ser ejecutadas conductualmente, pero la cuestión de qué o quién las ejecute es lo único que no importa o que es irrelevante en este contexto.

El estudio del lenguaje como ligado a la conducta tiene la ventaja de que permite generar modelos de procesamiento de información que pueden ser aplicados, *inter alia*, a humanos, animales, máquinas, cerebros, códigos genéticos, etc. Se puede, por ejemplo, construir un algoritmo y diseñar una máquina que lo ejecute, pero también se puede estudiar la actividad de un órgano u organismo para tratar de analizar su "conducta" en términos del mismo paradigma, es decir, el de procesamiento de información. La inteligencia artificial y la robótica están haciendo lo primero; las neurociencias, las ramas derivadas de la fisiología de lo computable, la neuroetología, la psicología e incluso la lingüística están haciendo lo segundo. Pero lo importante es notar que no hay la menor razón para aceptar que cualquiera de ellas estudia o se ocupa de la mente o de algo "mental": lo que estudian es sencillamente la relación entre la conducta y la clase de procesamiento de información que suponemos que se requiere para ejecutarla.

Si todas las ramas de las ciencias cognitivas pueden aglutinarse alrededor del paradigma de procesamiento de información es porque dicho paradigma, al no tener ninguna clase de compromiso con la clase de entidad que procesa la información, abre la posibilidad de atrapar el análisis de un procesamiento de información tanto en abstracto como ligado a una entidad concreta, puesto que es obvio que la conducta asociada con un programa puede involucrar o no una actividad física. La expresión "Suma los y dos" y la expresión "Trae el lápiz que está sobre la mesa" pueden ser identificadas simplemente como algoritmos que pueden ser ejecutados por cualquier tipo de entidad que esté capacitada para procesarlo o ejecutarlo conductualmente. Lo único que se asume es que el algoritmo (es decir, el lenguaje) expresa

actividades posibles tales que, si se ha construido la máquina adecuada para ello, a su vez puede desencadenar conductas específicas. Algunas de estas conductas exigen un movimiento corpóreo, otras no. Si a un tipo de conducta se le llama 'física' y a la otra 'mental', ello se debe a que no contamos con otro término para referirnos a la clase de conducta que no necesariamente implica un movimiento corpóreo. Pero ningún misterio está aquí implicado.

Es legítimo interpretar la conducta que llamamos 'mental' como algo que emerge de un algoritmo, siempre y cuando se comprenda que lo que emerge del algoritmo no es una entidad ni una actividad etérea, sino simplemente conducta. De hecho, el algoritmo se concibe como aquello que efectivamente puede generar una conducta, por lo que el mismo concepto de algoritmo acarrea consigo la idea de conducta, está lógicamente vinculado a ella. Por eso, un algoritmo puede ser objeto de experimentación: podemos probar lo que expresa un algoritmo por el efecto que produce como conducta. Pero suponer, como lo hacen muchos, que el algoritmo es la expresión de la actividad de un "algo", una mente, o que del algoritmo surge la entidad denominada 'mente', es tan sólo el resultado de una grave confusión. Ninguna mente está aquí involucrada: lo único que hay son signos, reglas, conducta. Ahora bien, parecería que no es así como los cognitivistas están tratando de reenfocar el problema mente-cuerpo. Empero, si lo que hemos dicho es correcto, entonces puede verse que los nuevos reduccionismos y emergentismos que se encuentran diseminados en todas las áreas de las ciencias cognitivas no son otra cosa que el resultado de una mala interpretación de sus propios conceptos y de los alcances de la metodología que utilizan.

Si vemos las cosas desde una perspectiva histórica podremos percatarnos de que el problema mente-cuerpo, en relación con el lenguaje y la conducta que puede ser expresada a través de o gracias a él, en el fondo ni se ha tocado. Lo que por nuestra

parte hemos tratado de hacer ver es simplemente que hay estrategias filosóficas que permiten abordar el estudio del pensamiento y de la conducta vía el estudio del lenguaje sin para ello vernos hundidos en el famoso problema mente-cuerpo. Este problema sigue pendiente y, por lo mismo, sigue siendo parte del quehacer filosófico. Nadie afirma que sea un error filosófico tratar de utilizar los nuevos resultados o avances de las ciencias para tratar de esclarecer un viejo problema filosófico. Lo que afirmamos es que la ciencia, al menos como la conocemos hasta ahora y por las razones que se han expuesto a lo largo de este trabajo, no tiene por qué abordar, para progresar, problemas que la misma filosofía aún considera como rompecabezas, esto es, como un conjunto de piezas sueltas que aún no sabemos siquiera cómo acomodar.

El problema de la relación mente cuerpo se terminó de imprimir en agosto de 1993, en Grupo Editorial Interlínea, S. A. de C. V., Irlanda 121 Bis, col. Parque San Andrés, Coyoacán, tel. 689 4829. Se imprimieron 1 000 ejemplares sobre papel cultural de 36 gr; forros en cartulina couché de 210 gr; la edición estuvo al cuidado de José Antonio Robles

METAFÍSICA, PROCESOS MENTALES Y OBSERVACIONES EMPÍRICAS

Nydia Lara Zavala
Francisco Cervantes Pérez y
Arturo Franco Muñoz

Laboratorio de Neurocomputación
Centro de Instrumentos, UNAM

INTRODUCCION

Este trabajo tiene como propósito cuestionar la idea de que es posible estudiar, a través de la observación de los patrones de actividad que manifiesta el sistema nervioso central (SNC), operaciones cognitivas de diversa índole. Lo que se quiere demostrar es que detrás de la propuesta misma de que el SNC es capaz de pensar, imaginar, reconocer, etc., hay una postura metafísica que guía la manera como se están interpretando los datos científicos. No es nuestra intención cuestionar los datos que se obtienen a través de los experimentos que estudian los diversos patrones de actividad que manifiestan diversos tipos de neuronas en diversas vías sensoriales cuando son expuestas a determinado tipo de estímulo. Lo que nos interesa revisar es el tipo de implicaciones que conlleva agregarle a esos mismos datos una determinada interpretación metafísica. Con esto queremos decir que no son los datos científicos, sino la elección de una particular postura metafísica lo que va a determinar la clase de características que se va a agregar a la materia que compone al SNC, para que sea plausible sostener que es por la manera como se activan los distintos componentes que lo conforman, que se observa cómo es que esos componentes son capaces de percibir y reconocer un determinado estímulo.

En este trabajo nos vamos a concentrar en dos de las posturas metafísicas más aceptadas por los investigadores que sostienen que existe una cosa tal como procesos mentales y que dichos procesos son explicables en términos de procesos cerebrales. La doctrina metafísica a las que nos referimos son, por un lado, la denominada 'teoría de la identidad' (Rosental, 1971, Flew, 1970) y, por el otro, la del emergentismo (Rudomín, 1978, Dennett, 1975, 1978^a, 1978b, 1983, Sperry, 1980, Bunge, 1977). Las dos teorías caen dentro de la posición filosófica conocida bajo el nombre de 'materialismo', pero mientras que los teóricos de la identidad consideran que es la materia misma la que se debe concebir como el agente que percibe, piensa, reconoce, etc., los emergentistas sostienen que no es la materia en sí la que posee dichas propiedades, sino que éstas emergen de su organización.

Cabe aclarar que ni los teóricos de la identidad ni los emergentistas aceptan que el SNC sea la causa de las actividades mentales. Para los teóricos de la identidad las actividades mentales son consideradas idénticas a los procesos cerebrales, en tanto que para

los emergentistas las actividades mentales surgen o aparecen como formas superiores de la organización de la materia. Dejando un poco de lado esta diferencia, lo que podemos afirmar es que tanto los teóricos de la identidad como los emergentistas coinciden en la suposición que sostiene que la descripción física de un mecanismo cerebral equivale a la descripción de una actividad mental.

Con la ayuda de las dos teorías, diversos investigadores se sienten dispuestos a aceptar sin cuestionar que:

- a) el cerebro tiene actividades mentales
- b) el cerebro tiene representaciones sensoriomotoras y
- c) el cerebro se forma una imagen interna de la realidad externa.

Para que esto tenga sentido, es importante aclarar dos puntos. El primero es el hecho de que ninguna de las dos teorías cuestiona la idea de que existe una cosa tal como procesos o actividades mentales y, el segundo es que, en lo único que difiere una teoría de la otra es sólo respecto a qué se debe considerar como el agente que percibe, piensa, imagina, etc.: la materia, en el caso del teórico de la identidad, o su organización en el caso del emergentista. En este sentido podemos afirmar que aunque estas doctrinas rechazan la idea de una mente que puede existir independientemente del cuerpo, no rechazan sus productos, por lo que en última instancia lo único que le cuestionan a un mentalista es la idea de que los procesos mentales sean el producto de una sustancia no material. Esto lo logran simplemente agregándole a la materia o a su organización las características o procesos que el mentalista ponía en la sustancia mental. Por lo que la discusión entre materialistas y mentalistas no es mas que una querrela sobre el tipo de agente que debe considerarse poseedor de las misteriosas actividades mentales.

Asimismo, vale la pena resaltar que, aunque los teóricos de la identidad no reconocen la distinción entre la mente y el cuerpo, los emergentistas enérgicamente rechazan la idea de que la materia sea capaz de pensar, por lo que se puede decir que ésta última forma de materialismo sí acepta la diferencia entre lo mental y lo corpóreo, aunque hace depender a lo mental de lo corpóreo. Para el emergentista no es la materia en sí misma la que produce los procesos mentales, como propondría un teórico de la identidad, sino que es de las formas "superiores" de la organización de la materia en movimiento de donde emerge la conciencia. Independientemente de los problemas que cada una de estas teorías acarrea, en los dos casos está implícita la idea de que si somos capaces de conocer los procesos cerebrales, entonces, en principio, somos capaces de conocer los procesos mentales. De aquí se deriva la idea de que es posible utilizar el término 'cerebro' como un término sinónimo del de 'mente', o, al menos, en más de una ocasión se sostiene que ciertas operaciones normalmente asociadas con el uso del término 'mente' (e.g., percibir, iniciar movimientos voluntarios, tener sensaciones, formarse imágenes mentales, etc.) se expresan sustituyendo el término 'mente' por el de 'cerebro'. Esto no sería un problema si el intercambio de términos fuera netamente lingüístico, es decir, si los que utilizan estos términos como sinónimos estuvieran dispuestos a aceptar que es exactamente lo mismo decir que "el cerebro percibe" a decir que "la mente percibe". Todos podemos utilizar un término como mejor nos parezca, pero graves inconsistencias empiezan a aparecer cuando hablamos del

conocimiento de la actividad cerebral en contraste con el conocimiento de la actividad mental.

La gran mayoría de los materialistas, en este caso, parece que no le dan mucha importancia al hecho de que, en principio, uno puede tener acceso cognitivo a sus propias actividades mentales (e.g., percepciones, deseos, intenciones, etc.) sin tener ningún tipo de conocimiento sobre cómo funciona el cerebro. Con esto queremos decir que, si los términos fueran, conceptualmente hablando, intercambiables, el análisis introspectivo de las propias actividades mentales, en principio, podría llevarnos al conocimiento de las actividades cerebrales, sin la necesidad de recurrir a ningún tipo de actividad empírica para conocer el cerebro, su organización y su funcionamiento; cosa que, por supuesto, es falsa.

Los materialistas, sin embargo, podrían argüir que lo que ellos afirman es que es el análisis empírico del funcionamiento completo del sistema nervioso lo que, en principio, puede llevarnos al conocimiento de todas las actividades mentales (Romo, 1990) puesto que lo que ellos sostienen es que el cerebro, en el caso de los teóricos de la identidad, o la compleja organización de sus actividades, en el caso de los emergentistas, es lo que lleva a cabo las actividades mentales. En este sentido, tanto los teóricos de la identidad, como los emergentistas, parece que piensan que la descripción del funcionamiento de un mecanismo cerebral equivale a la descripción de los procesos mentales que lleva a cabo el cerebro.

La pregunta que puede plantearseles a esta clase de investigador es si realmente cree que lo que se está postulando como "actividades mentales" (e.g., percepción, representación mental, acción voluntaria, etc.) son elementos o mecanismos observables en el mismo sentido en que el mecanismo de un reloj, o la actividad de una neurona, o de grupos de neuronas es observable. Georgopoulos y col. (1989) sostiene que, si bien "una visualización directa de una operación cognitiva en términos de la actividad neuronal en el cerebro aún falta" ("a direct visualization of a cognitive operation in terms of neural activation in the brain is lacking" (Science. Vol. 243, 13 January 1989, p. 234) es posible solucionar este "problema" analizando la operación cognitiva que se realiza cuando se observa la rotación mental imaginaria que llevan a cabo las neuronas un poco antes de que se produzca un movimiento voluntario.

Ahora bien, es evidente que estos investigadores caen en una contradicción al afirmar que, aunque no ha sido posible observar una operación cognitiva en términos de la actividad cerebral, ellos pueden resolver este problema observando una operación cognitiva (i.e., la rotación mental imaginaria) antes de que las neuronas lleven a cabo otra operación cognitiva (i.e., un movimiento voluntario).

El reconocimiento de este desacuerdo conceptual nos lleva a pensar que podría ser interesante revisar otros trabajos ligados a las neurociencias en donde prestigiosos investigadores sostienen que observando la actividad de las neuronas es posible descubrir los mecanismos de las operaciones cognitivas que llevan a cabo esas neuronas. Si consideramos que esto es un equívoco conceptual, lo que parece conveniente es tratar de establecer exactamente qué es lo que observan estos investigadores y qué es lo que dicen que

observan. Para este propósito revisamos varios trabajos del área de la neurofisiología, tomando como punto central el trabajo de Phillips y col. (1988), titulado "Spatial pattern representations and transformation in monkey somatosensory cortex". Lo que vamos a demostrar en este trabajo es que en ningún caso es posible observar en la actividad cerebral algo más que la actividad cerebral misma y que el supuesto de operaciones cognitivas involucradas en dicha actividad no es demostrable empíricamente.

LA PRETENSION DE ESTUDIAR PERCEPCION, MEMORIA Y APRENDIZAJE A TRAVES DE LA ACTIVIDAD CEREBRAL

En el área de las neurociencias frecuentemente se plantea, como objetivo principal, el estudio de los posibles mecanismos neuronales que conciernen a la transformación y representación de las características espacio temporales de lo que ellos llaman 'procesamiento de información' sensorial que se lleva a cabo dentro del SNC. Lo que no se considera es que el sólo hecho de hablar de 'procesamiento de información', ya presupone la aceptación de lo que filosóficamente se conoce con el nombre de la 'teoría causal de la percepción' propuesta por pensadores como Descartes, Locke y Boyle desde el siglo XVII. Dicha teoría debe ser entendida como la responsable de introducir en el pensamiento moderno no sólo la distinción entre procesos mentales y procesos físicos, sino la gama de discusiones en torno al agente que se supone percibe, imagina, desea, piensa, etc. De estas discusiones surgen diversas teorías mentalistas al igual que diversas teorías materialistas, de las cuales nosotros sólo hemos abordado dos casos: la de la identidad y la del emergentismo, por considerarlas las más populares entre los neurofisiólogos.

Ahora bien, nosotros mencionamos más arriba que ni el teórico de la identidad, ni el emergentista, aceptan que el SNC sea la causa de los procesos mentales, pero esto no los lleva a rechazar lo que se conoce como el 'dualismo ontológico'. Lo que sostiene dicho dualismo es que el origen de nuestras percepciones parece requerir de la existencia de una realidad externa y de un sujeto que la capte, la procese y la convierta en una imagen mental que le sirva como la representación interna de la realidad externa. Bien entendido, el problema que plantea este dualismo no consiste en suponer que la percepción sensible es un producto de la mente, sino en el hecho de que aceptado el dualismo entre realidad externa y representación interna, la percepción queda entendida como un algo de carácter mental pero dependiente de la actividad del SNC del sujeto. Lo que discuten los mentalistas y los materialistas es si las representaciones mentales a que dan lugar los procesos involucrados en éste tránsito, deben ser consideradas como la actividad de una sustancia no física o como la actividad de la sustancia física que compone al cerebro. Si asumimos lo segundo, lo que obtenemos en términos gráficos es la imagen que se presenta en la figura 1.

Asumido el dualismo ontológico, la teoría causal de la percepción lo que diría, utilizando un lenguaje materialista, es que una de las funciones de los sistemas sensoriales del SNC refiere a la posibilidad de transformar la representación primaria que provoca un estímulo - que se genera por medio de los mecanismos sensoriales periféricos - en procesos

superiores que desembocan en percepciones del estímulo. Evidentemente, una investigación empírica que se guía bajo lo que *a priori* esta asumido por la teoría, no nos puede descubrir nada acerca de la percepción, ni de ningún otro "proceso mental". Lo que se hace es simplemente darle contenido físico a lo que la teoría metafísica *a priori* estipula como posible interpretación de los datos que justifican la idea de que el cerebro percibe, piensa, planea, aprende, memoriza, etc. No obstante, últimamente se han desarrollado infinidad de investigaciones utilizando diferentes vías sensoriales (i.e., visual, táctil, olfativo, etc.) (Phillips y col., 1988; Livingstone, 1988; Mishkin y Appenzeller, 1987; Ewert, 1980) supuestamente para "descubrir" los componentes e interacciones que la teoría causal de la percepción previamente ya había estipulado como necesarios para darle contenido empírico a lo que la teoría sostiene.

Un ejemplo claro de esta aproximación, y que se encuentra con pequeñas variables en gran parte de los trabajos de neurociencias, es el artículo de Phillips y col. (1988) "Spatial pattern representation and transformation in monkey somatosensory cortex" donde se investiga el procesamiento de información táctil. Sobre éste tema, lo que sostiene la teoría causal de la percepción es que el cerebro esta ligado a receptores que se encuentran distribuidos por toda la piel. Algunos de esos receptores responden a la presión al entrar en contacto con un objeto, otros responden al calor y otros al frío. Otros receptores de la piel y del cuerpo responden a una gran variedad de estímulos transmitiéndole al cerebro impulsos que finalmente causan sensaciones de dolor o placer. Hay además otros sentidos, como son los cinéticos, donde se supone la presencia de receptores en los músculos, los cuales también mandan impulsos nerviosos al cerebro donde se detecta la posición de los miembros, los cuales inconscientemente se ajustan para guiar o llevar a cabo un movimiento voluntario (Hirts, R.J., *Perception*, Encyclopedia of Philosophy, Vol. 6, Macmillan Publishing Co., Inc. & The Free Press, N.Y., 1967.)

Siguiendo esta forma de ver las cosas, lo que hacen Phillips y col. en su trabajo es registrar, en el mono *Macaca mulatta*, los potenciales de acción de las fibras aferentes primarias y el de la corteza somatosensorial conforme rotan un tambor con una letra en alto relieve sobre la parte de la yema de un dedo de la mano, que ellos identifican como el campo receptivo de ambos tipos de neuronas. Estos registros les permiten caracterizar las distintas propiedades espacio-temporales de las respuestas de estas neuronas ante el mismo estímulo. La localización del tambor sobre el campo receptivo y los tiempos donde se producen potenciales de acción son monitorizados continuamente por una computadora que traduce cada potencial de acción como si fuera un punto. Una vez que se pasa todo el tambor sobre el campo receptivo de la neurona, se le pide a la computadora que reúna, en una sola imagen, todos los potenciales de acción obtenidos en diferentes tiempos, traduciendo cada potencial como si fuera un punto. Con esto conforman lo que ellos definen como un SEP ("Spatial Event Plot"). Los SEPs que se obtienen a partir de los registros de las aferentes primarias se muestra en la fig. 2; mientras que los que se obtienen de los registros del área 3b y del área 1 de la corteza se muestra en las figuras 3 y 4.

Lo que estos autores encuentran es que, en algunos registros es posible traducir la manera como se producen los potenciales de acción, como si estos materialmente copiaran

la forma del estímulo cuando ellos forman las SEPs (ver fig. 5), esto es, cuando reúnen todos los potenciales como si fueran una imagen hecha de puntos. Ante estos resultados Phillips y col. proponen dos tipos de interpretaciones: la conservadora y la suya. La primera diría simplemente que los SEPs son la caracterización de las propiedades espacio-temporales de la respuesta de una neurona provocada por un estímulo complejo en movimiento; mientras que la de ellos diría que el Sep de una sola neurona es sólo una aproximación espacio-temporal de la imagen neuronal que puede obtenerse de un grupo de neuronas, si suponemos que:

- a) existe un grupo numeroso de poblaciones de neuronas que manifiestan propiedades de respuesta semejantes a la que observamos en una sola neurona;
- b) sus respuestas localizadas están más o menos distribuidas de manera uniforme en toda la superficie de la piel; y
- c) sus campos receptivos están distribuidos sobre la piel con suficiente densidad como para producir una imagen espacio-temporal completa del estímulo.

Según éstos autores, si todo esto es factible, en principio, nosotros podríamos observar en las neuronas la imagen mental que produce en ellas un objeto externo; o para ponerlo en sus propios términos, podríamos observar la manera como el cerebro percibe éstas imágenes y las representa mentalmente. Esto es, según ellos, si se puede producir la imagen isomórfica del estímulo simplemente registrando la actividad de una población de neuronas, lo que podríamos observar es la manera como las células del cerebro perciben y representan al estímulo.

Esta forma de aproximar el problema también la encontramos al revisar trabajos en los que se describe el estudio de los posibles mecanismos neuronales que utiliza el SNC de los seres vivos durante el denominado 'procesamiento visual'. Por ejemplo, Livingstone (1988) sostiene que el cerebro puede identificar e interpretar las imágenes que él mismo es capaz de crear a partir de información sensorial detectada por la vista. El tipo de experimento que lleva a cabo refiere a registros electrofisiológicos de la actividad de las células de cada una de las diferentes subdivisiones del sistema visual (las vías blob, parvo-interblob y magno de las áreas visuales 1 y 2), tratando de averiguar a cuáles características del estímulo responden mejor las distintas células de cada una de éstas subdivisiones. Según Livingstone las células contenidas en los blobs son altamente selectivas al color o a la brillantez pero son indiferentes a la forma y al movimiento; mientras que las células de las regiones interblob son selectivas a la orientación del estímulo pero no al color o al movimiento. Según su propio ejemplo, una célula interblob puede responder a una barra vertical independientemente de como se mueva la barra o de si es negra, blanca o coloreada; en sus términos, el único "criterio" que utiliza la célula es que la barra sea vertical; la misma célula, de acuerdo a sus propios registros, no responde a barras con otra orientación. De manera semejante, Livingstone a través de registros de otro tipo de células, observa que éstas se activan o bien ante una barra horizontal que se mueve hacia arriba, o bien ante una barra vertical que se mueve horizontalmente, pero la misma célula no responde a ambas barras. Este resultado lo llevo a concluir que existen células que no son selectivas al color pero que son selectivas a la orientación y al movimiento.

Esta autora, al describir el denominado 'procesamiento de información' que supone se lleva a cabo en la vía visual primaria y en la corteza visual de los mamíferos, postula que la organización de las regiones del cerebro involucradas está constituida al menos por tres sistemas separados de procesamiento de información: la vía blob, que es la parte del sistema visual que reconoce la forma de los objetos; la vía parvo-interblob, que procesa información para la percepción del color; y la vía magno que se encarga de procesar información para la percepción del movimiento y las pistas que utilizan para juzgar la distancia a la que se encuentran los objetos, así como su localización y organización espacial (ver figura 6).

De manera semejante al trabajo de Livingstone, Mishkin y Appenzeller (1987), postulan que la información visual es procesada de manera secuencial a lo largo de una vía del SNC compuesta por diferentes estructuras neuronales (retina, cuerpo geniculado lateral, distintas áreas de la corteza visual y corteza temporal inferior) y que, cuando se analizan las propiedades tanto anatómicas como fisiológicas de los elementos neuronales de estas estructuras, se observa que conforme se avanza hacia las estructuras más internas de la vía visual, no solamente se incrementan las dimensiones de los campos receptivos de las neuronas ("su ventana hacia el mundo visual"), sino que también se enriquece la complejidad del procesamiento que pueden llevar a cabo las distintas células con base en la información que reciben; esto es, una neurona de una estructura intermedia responde acorde al procesamiento que lleva a cabo de la información que resulta del procesamiento realizado por un grupo de neuronas de la estructura inmediata anterior. Estos autores, al igual que Phillips y col. y Livingstone, sostienen que distintas células responden selectivamente a más de una de las propiedades físicas de los objetos (incluyendo tamaño, figura, color y textura), hasta que en la estación final de la corteza temporal inferior, las células sintetizan una completa representación de los objetos externos (ver figura 7).

Por otro lado Mishkin y Appenzeller (1987), dicen que "a lo largo de la vía visual el cerebro integra datos sensoriales en experiencias perceptivas" y que ellos investigan la participación de ciertas estructuras profundas del cerebro y sus interacciones con vías perceptivas de la corteza visual para dar cuenta de la transformación de estímulos sensoriales en recuerdos. Es decir, estos autores están buscando los mecanismos cerebrales que dan lugar al recuerdo. Con base en experimentos de lesiones cerebrales, proponen la existencia de dos circuitos amplios, uno ligado a la amígdala y otro al hipocampo, los cuales, según ellos, son responsables de muchos tipos de aprendizaje cognitivo (*i.e.* la capacidad de reconocer un objeto familiar, recordar sus cualidades sensoriales no percibidas, recordar su localización anterior y el significado emocional asociado al objeto). Pero lo que queremos resaltar de este trabajo es que éstos autores amplían la descripción dada por Livingstone, sobre lo que las neuronas son capaces de percibir, aparentemente con el objeto de incluir otro tipo de capacidades mentales del cerebro, como son la memoria y el aprendizaje.

DISCUSION

Aunque el trabajo de Phillips y col. (1988), el de Livingstone (1988), y el de Mishkin y Appenzeller (1987), estudian vías sensoriales distintas, hay en ellos una estructura común

en la manera de interpretar sus datos, aunque difiera la manera como tienden a multiplicar a los percipientes que según ellos habitan en los cerebros.

Para Phillips y col. , por ejemplo, cada neurona percibe parcialmente la información de un estímulo y la inferencia de éstos investigadores es que si pudiésemos obtener el registro de la percepción individual de una población de neuronas podríamos observar una imagen nítida de lo que percibe el cerebro. Para Livingstone y Mishkin y Appenzeller no sólo tenemos la percepción parcial de cada neurona, sino que además tenemos vías completas que a su vez sólo perciben parcialmente la información de un estímulo. Cada vía sensorial, en el caso de éstos autores, se encargan de integrar la información parcial que cada neurona percibe del estímulo, pero la integración de la percepción individual de una población de neuronas de una sola vía, sólo nos da información parcial de las características de un estímulo. En éste caso, se dice que es la integración de lo que percibe cada vía en algunas estructuras cerebrales, lo que en última instancia podría darnos la imagen de lo que percibe o recuerda el cerebro.

Sin embargo, si comparamos el trabajo de Phillips con el de Livingstone y el de Mishkin y Appenzeller podemos encontrar que aunque éstos autores estudien distintas vías sensoriales, y distintos procesos "mentales", la diferencia se encuentra sólo en la complejidad de la forma como ellos creen que se procesa o almacena la información, pero no en la manera como fundamentan que el cerebro percibe.

Lo que todos éstos autores sostienen es que cada célula percibe una parte del estímulo que al unirse con la percepción de las otras células logra que el cerebro perciba la imagen nítida del estímulo. Esto parece implicar que, para ellos, cada célula puede ser interpretada como si fuera un percipiente, aunque sea parcial y que, la percepción unificada de todas las células, dan como resultado lo que percibe el cerebro. Su interpretación involucra, en todo caso, la aceptación de más de un percipiente: cada célula, cada estructura, cada vía cerebral, más el cerebro, como si el cerebro fuera algo más que el conjunto de sus células o estructuras.

No obstante, el punto que nos interesa destacar es que lo que provoca ésta forma de interpretar los datos experimentales nos lleva del discurso metafísico al científico y del científico de nuevo al metafísico. Pero ahora encontramos todo mal ubicado. Decimos esto porque, bien entendido, los datos experimentales en ningún caso pueden confirmar si las neuronas, las estructuras, las vías cerebrales o el cerebro perciben; más bien sólo sirven para demostrar qué neuronas, qué estructuras o qué vías se activan cuando exponemos al animal completo a cierto tipo de estimulación (Cervantes-Pérez y Lara, 1991). Además, cuando intentamos definir a qué corriente filosófica pertenecen éste tipo de interpretaciones, topamos con que, salvo el discurso que se deriva de la teoría causal de la percepción, no queda claro si éstos autores piensan en términos de un emergentista, o en términos de un teórico de la identidad. Como emergentistas podríamos suponer que la neurona tiene la organización suficiente para percibir, pero esto es poco menos que increíble. Como teóricos de la identidad, podríamos suponer que la misma actividad que se observa en la neurona equivale a la observación de la actividad mental de la neurona, lo cual sin duda introduce

gratuitamente la gama de conflictos que caracteriza al problema mente-cuerpo en todas y cada una de las neuronas. La impresión que todo esto genera es que lo que hacen todos estos investigadores es una especie de mezcla de la teoría de la identidad con la del emergentismo, como si se tratara de una única teoría. Se toma un poco de una y un poco de otra y se olvidan que todo no sólo resulta sumamente confuso, sino que como cada teoría tiene principios que no necesariamente son los mismos que acepta otra teoría, el eclecticismo resulta inválido, tanto epistemológica como estructuralmente. Esto es claro sobre todo cuando reparamos en el hecho de que, lo que observan éstos autores, no va más allá de una correlación entre el patrón de estimulación y el patrón que puede observarse de la actividad de ciertas neuronas estimuladas por medio de las fibras aferentes y los receptores, los cutáneos en el caso de Phillips y col. y los fotoreceptores en el caso de Livingtone y Mishkin y Appenzeller.

No obstante, lo que ellos suponen que observan es la manera como las neuronas están representando mentalmente al estímulo; por lo que, según ellos, cualquier investigador es capaz de observar, no sólo a la neurona y su actividad, sino que también cómo selecciona, representa, percibe y recuerda ésa neurona o grupos de neuronas a un estímulo específico. Quizá un ejemplo, extraído del trabajo de Phillips y col., puede aclarar por qué esta idea puede considerarse absurda.

Cuando se forma un SEP, siguiendo la técnica de Phillips y col., todo parece indicar que lo que confunde a éstos investigadores, es que ellos, en efecto, lo que perciben es una doble imagen: la del estímulo que se utiliza para activar a los receptores y la que se gesta cuando todos los potenciales de acción de las células de las estructuras corticales o subcorticales, según sea el caso, son puestos juntos como puntos distribuidos en el espacio formando una sola imagen. Su interpretación de ésta doble imagen los lleva a sostener que, la que pertenece al estímulo es la imagen real y la que se obtiene del registro es la imagen percibida por las neuronas de la corteza. Claro que sus registros les indican que existen neuronas que se activan de cierta manera ante determinadas características del estímulo y se activan de otra o no se activan, ante otra característica del estímulo. Lo que ellos interpretan de ésta variedad de registros es que hay neuronas cuya tarea consiste sólo en detectar una característica específica del estímulo, como si dentro del SNC existiera una especie de organización social estructurada, la cual según su interpretación se expresa como una división del trabajo en donde cada una de las neuronas se dedica selectivamente a percibir sólo una parte de la realidad externa. Por supuesto que como, según ellos, lo que percibe una sola neurona no es suficiente como para obtener una imagen nítida de la característica que le corresponde obtener del estímulo (e.g., forma, tamaño, textura, orientación, color, etc.), suponen que se requiere de la participación de un grupo numeroso de neuronas para lograrlo. Así, suponen que la labor conjunta del trabajo de las neuronas es suficiente para que el cerebro logre obtener una representación mental de las distintas características del estímulo.

El problema aquí es que, cuando el observador percibe el patrón de actividad de una neurona en correlación con alguna característica del estímulo, piensan que la neurona se activa para percibir al estímulo. Ellos, como dijimos al principio, están suponiendo que

estudian los mecanismos a través de los cuáles se gestan las representaciones en el cerebro; sin embargo, lo que hacen se reduce a registrar la actividad de una neurona cuando el animal es expuesto a determinado tipo de estimulación. Pero si nosotros pensamos, por ejemplo, que en lugar de una neurona, tenemos una fotocelda sintonizada a activarse sólo cuando la exponemos, digamos, a determinada intensidad de luz azul y tenemos un estímulo que es un tambor con una letra k iluminada con una lámpara de luz azul; es obvio que al presentar el estímulo sobre el campo receptivo de la fotocelda y monitoriar el patrón de actividad de esa fotocelda hasta poder formar un SEP, sin duda obtendríamos, utilizando la metodología de Phillips y col., una representación isomórfica del estímulo; pero ¿podríamos concluir de aquí que lo que estamos registrando es la manera cómo ésa fotocelda percibe el estímulo?

Si seguimos la tendencia explicativa de estos investigadores, cualquier tipo de materia (biológica o física) se vuelve perceptiva y, en ese sentido, si seguimos su interpretación, seguramente lo que según ellos estamos observando es la manera como la fotocelda percibe el estímulo. Esto es, si nos basamos en la evidencia de la obtención de un registro que correlacione alguna característica de un estímulo dado con la actividad específica de diversas neuronas o de fotoceldas sintonizadas o estructuradas para activarse sólo a cierto tipo de estimulación, ya no importa que sea una neurona, una fotocelda, o el mismo instrumental que nos permite registrar la actividad de una neurona o de una fotocelda (e.g., amplificadores, osciloscopios, polígrafos, etc.), todo se podría interpretar como si registraríamos la capacidad de éstos elementos para percibir.

Suponemos que ningún investigador seriamente aceptaría la posibilidad que el registro de la activación de una fotocelda signifique el registro de la percepción de la fotocelda. Esta negación debería de hacernos recapacitar sobre lo que estos investigadores están tratando de sostener apoyados, no en sus registros, sino en la interpretación que están haciendo de ellos. Sin embargo, no es fácil corregir creencias metafísicas y esto es lo que ha generado en los últimos años, no sólo la sorprendente discusión sobre la posibilidad o imposibilidad de construir máquinas computacionales con capacidad de percibir, conocer o sentir (Searle, 1990; Churchland y Smith Churchland, 1990), sino la extraña idea que sostiene la necesidad de plantear un "nuevo" dualismo: el que separa a la materia biológica pensante del resto de la materia (Cervantes Perez y Lara, 1991).

Ahora bien, si uno es capaz de distinguir entre lo que genuinamente hacen éstos autores de lo que dicen (o creen) que hacen, parece que podemos encontrar problemas desde la manera como exponen su hipótesis de trabajo. Lo que ellos creen que están tratando de analizar es cómo una función de los sistemas sensoriales puede transformar una representación primaria de un estímulo en una percepción, por lo que de origen ya hay una predisposición a suponer que la neurona percibe, lo cuál culmina en la aceptación de que cualquier cosa cuyo patrón de actividad forme una especie de SEP, percibe. Lo que hacen estos autores en realidad se limita a comparar la actividad de las aferentes primarias respecto de la actividad de las neuronas de la corteza ante un mismo estímulo. Lo que encuentran es que el patrón de actividad de algunas neuronas ante el estímulo se asemeja, mientras que la de otras es distinta. Estos registros sólo sugieren que existen neuronas que se activan de

cierta manera ante determinado tipo de estímulo y se activan de otra manera, o simplemente no se activan, ante otro tipo de estímulos. Sin embargo, lo que ellos sostienen es que hay neuronas que se dedican a percibir determinada característica de cierto tipo de estímulo, mientras que otras se dedican a percibir otras características diferentes del mismo estímulo, como si las neuronas pudieran "decidir" qué es lo que les corresponde percibir y qué no les corresponde percibir. Pero este animismo cerebral podría evitarse si damos otra interpretación a los mismos datos.

Aunque suene menos atractiva que las interpretaciones metafísicas que se nos ofrecen en estos trabajos, podemos decir que lo que éstos autores encuentran es que existen mecanismos específicos en el SNC para responder sólo a cierto tipo de estimulación. Esto, en sí mismo, es más que importante que lo que ellos interpretan porque, al conocer la especificidad de los mecanismos del SNC, podemos empezar a entender no sólo cómo funcionan, sino también cómo podemos repararlos si somos capaces de modificar, de alguna manera, sus respuestas dinámicas (Bermúdez y cols., 1987). Si despejamos el hallazgo de éstos autores de todo su lenguaje mentalista, lo que podemos ver es que, mientras ellos creen que están estudiando percepción, lo que genuinamente nos están ofreciendo es una herramienta para buscar formas de control, manipulación, corrección e incluso de reparación del SNC, así como la posibilidad de construir máquinas y prótesis artificiales. Sin embargo la importancia de lo que encuentran éstos autores se diluye al confundir el estudio de los receptores con el estudio de la percepción, lo cuál hace de su trabajo, sin duda una labor interesante, pero, en cierta medida, irrelevante en lo que refiere al potencial de investigación que los mismos artículos conllevan.

AGRADECIMIENTOS. Este trabajo fue financiado parcialmente por la Dirección General de Asuntos del Personal Académico (DGAPA) de la UNAM a través de los donativos IN 026189.

BIBLIOGRAFIA

- AMSTRONG, D. M. , The nature of mind, Inaugural lecture of the Chllis professor of Philosophy at the University of Sydney (1975), aparece en *The Mind Brain Identity Theory*, ed. por A. G. N. Flew, St. Martin s Press, N.Y. , 1970.
- ARBIB, M. , *Brain, Machines and Mathematics*, Springer Verlag, Berlin, N. Y. , 1987.
- BARR, A. & FEIGENBAUM, E. A. , *The Hand Book of Artificial Intelligence*, William Kaufman, Inc., Los Altos, California, 1981.
- BERMUDEZ, F. , FERNANDEZ, J. , SANCHEZ, M. A. , AGUILAR, R. and DRUCKER COLIN, R. , Fetal brain transplant induce recuperation of taste aversion learning, *Brain Research*, num.416, p. 147 152, 1987
- BUNGE, M. , Emergence and the Mind, *Neuroscience*, Vol. 2, p. 501 510, Pergamon Press, Great Britain, 1977
- CHURCHLAND, P., Scientific realism and the plasticity of mind, Cambridge Univ. Press, Cambridge, 1979.

- Matter and consciousness: A contemporary introduction to the philosophy of mind**, MIT Press, Cambridge, Mass, 1984.
- Reduction, qualia, and the direct introspection of brain states**, *Journal of Philosophy*, Vol. 82, p. 8-28, 1985.
- CHURCHLAND, P. Y SMITH-CHURCHLAND, P.**, Could a machine think?, *Artificial Intelligence: a debate*, *Scientific American*, enero 1990.
- CORNMAN, J. W.**, The Identity of Mind and Body, *The Journal of Philosophy*, LIX, 18, p. 486-492, 1962; *aparece en Materialism and the Mind-Body Problem*, ed. David M. Rosenthal, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1971.
- DENNET, D. C.**, Brain writing and mind reading, *aparece en Language, mind and knowledge*, ed. K. Gunderson, *Minnesota Studies in the Philosophy of Science* 7, Univ. of Minnesota Press, Minneapolis, 1975.
- Artificial intelligence as philosophy and as psychology**, *Philosophical perspective on artificial intelligence*, ed. M. Ringle, Humanities Press & Harvester Press, N. Y., 1978a.
- Brainstorms: Philosophical essays on mind and psychology**, MIT Press, Cambridge, Mass., 1978b.
- Intentional systems in cognitive ethology: The "Panglossian Paradigm" defended**, *Behavioral and Brain Sciences*, Vol. 6, p. 343-390, 1983.
- EWERT, J. P.**, *Neuroethology. An introduction to the neurophysiological fundamentals of behavior*, Springer, Berlin, Heidelberg, N. Y., 1980.
- FEYERABEND, P. K.**, How to be a good empiricist: A plea for tolerance in matters epistemological, *Philosophy of science, the Delaware seminar*, Vol 2, ed., B. Baumrin, *Interscience* p. 3-39, N. Y., 1963a.
- Materialism and the mind-body problem**, *The Review of Metaphysics*, Vol. 17, p. 49-66, 1963b.
- FLEW, A. G. N.**, *The Mind/Brain Identity Theory*, *Antologia*, St. Martin's Press, N. Y., 1970.
- FODOR, J.**, *The Language of Thought*, Crowell, New York, 1975.
- Propositional Attitudes, Readings in Philosophy of Psychology**, Vol. 1, Ned Block, Harvard Univ. Press, USA, 1980.
- GEORGOPOULOS, A. P., LURITO, J. T., PETRIDES, M., SCHWARTZ, A. B., MASSEY, J. T.**, Mental Rotation of the Neuronal Population Vector, *Science*, Vol 243, p. 234-236, 1989.
- HUBEL, D. H., WIESEL, T. N.**, Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat, *Journal of Neurophysiology*, Vol. 28, p. 229-289, 1965.
- Receptive fields and functional architecture of monkey striate cortex**, *Journal of Physiology*, Vol. 195, p. 215-243, 1968.
- Functional architecture of macaque visual cortex**, *Proceedings of the Royal Society of London, Series B* 198, p. 1-59, 1977a.
- Plasticity of ocular dominance columns in monkey striate cortex**, *Philosophical Transactions of the Royal Society of London* 278/961, p. 377-409, 1977b.

- KORNBLITH, H.**, *Naturalizing Epistemology*, Antología, MIT Press, Cambridge, Mass., 1988.
- LARA, N.**, *La dinámica de la estructura de la percepción como herramienta cognoscitiva común de la filosofía y la ciencia*, Tesis de maestría en Investigación Biomédica Básica, UNAM, 1990.
- LIVINGSTON, M.**, *Art, illusion and the visual system*, Scientific American, Enero, 1988.
- MISHKIN, M.**, y **APPENZELLER, T.**, *The anatomy of memory*, Scientific American, 1987.
- MOUNTCASTLE, V.**, *Modality and topographic properties of single neurons of cat's somatic sensory cortex*, Journal of Neurophysiology, Vol. 20, p. 408-434, 1957.
- An organizing principle for cerebral function: The unit module and the distributed system**, ed. F. O. Schmitt & F. G. Worden; *The mindful brain*, MIT Press, Cambridge, Mass., 1979.
- PHILLIPS, J. R.**, **JOHNSON, K. O.**, **HSIAO, S. S.**, *Spatial pattern representations and transformation in monkey somatosensory cortex*, Proc. Natl. Acad. Sci., USA, Neurobiology, Vol. 85, p. 1317-1321, 1988.
- PLACE, U. T.**, *Is consciousness a brain process?*, aparece en la antología de A. G. N. Flew, *The Mind/Brain Identity Theory*, St. Martin's Press, N. Y., 1970a.
- Materialism as a scientific hypothesis**, aparece en la antología de A. G. N. Flew, *The Mind/Brain Identity Theory*, St. Martin's Press, N. Y., 1970b.
- PUTNAM, H.**, *Minds and Machines*, Dimensions of Mind: A Symposium, Sidney Hook ed., New York Univ. Press, N. Y., 1960.
- PYLYSHYN, Z.**, *What the mind's eye tells the mind's brain: A critique of mental imagery*, Psychological Bulletin, 80, 1-24, 1973.
- Imagery and Artificial Intelligence**, Readings in Philosophy of Psychology, Vol. 1, Ned Block, Harvard Univ. Press, USA, 1980.
- QUINE, W. V. O.**, *Epistemology Naturalized, Ontological Relativity and other Essays*, Colombia Univ. Press, N. Y., 1969.
- Natural Kinds**, aparece en *Naturalizing Epistemology*, Antología, MIT Press, Cambridge, Mass., 1988.
- ROMO, R.**, **ROMO, A.**, *La reconstrucción de la realidad*, Ciencias Cognitivas, ICyT, Vol. 12, No. 163, p. 29-36, 1990.
- ROSENTHAL, D. M.**, *Materialism and the Mind-Body Problem*, Antología, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1971.
- RUDOMIN, P.**, *El sistema nervioso: Maravilla que empezamos a descifrar*, Ciencia y Desarrollo, No. 21, 1978.
- SCHWARTZ, Computational Neuroscience**, MIT Press, Boston, 1990.
- SEARLE, J.R.**, *Is the brain's mind a computer program?*, Artificial Intelligence: a debate, Scientific American, enero 1990.
- SMART, J. J. C.**, *Sensations and brain processes*, aparece en la antología de A. G. N. Flew, *The Mind/Brain Identity Theory*, St. Martin's Press, N. Y., 1970.
- Further remarks on sensations and brain processes**, aparece en la antología de **FLEW, A. G. N.**, *The Mind/Brain Identity Theory*, St. Martin's Press, N. Y., 1970.
- SMITH CHURCHLAND, P.**, *Mind-brain reduction: New light from the philosophy of Science*, Neuroscience 7/5, p. 1041-1047, 1982.

The significance of neuroscience for Philosophy, TINS, Vol. 11, No. 7,
1988.

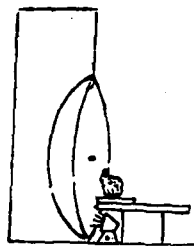
Neurophilosophy, MIT Press, Cambridge, Mass., 1989.

SPERRY, R. W., Mind-brain interaction: Mentalism, yes; dualism, no, Neuroscience, Vol 5,
p. 195-206, 1980.

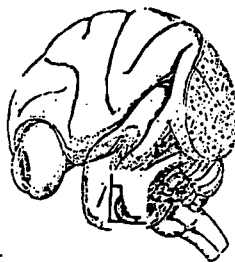
TOMASINI, A., Materialismo Dialéctico, Instituto de Investigaciones Filosóficas, UNAM,
comunicación personal, 1991.

WIENER, N., Cybernetics: Or control and communications in the animal and the machines,
Technology Press, Wiley, N. Y., 1948.

YOLTON, J., Biological Matter: Locke and French Medical Doctors in the Eighteen
Century, Rutgers Univ., comunicación personal, 1990



Realidad externa



Percepción de la realidad externa

Figura 1. La concepción del materialista metafísico respecto a cómo genera el cerebro la representación mental de la realidad externa.

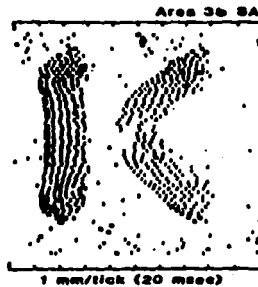


Figura 2. Los SEPs que se obtienen a partir de los registros de las aferentes primarias.

Peripheral SA, RA, PC

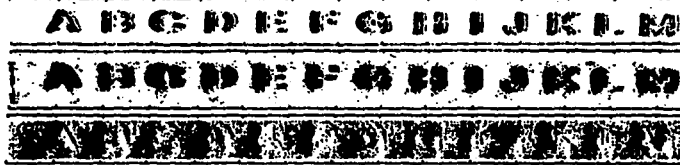
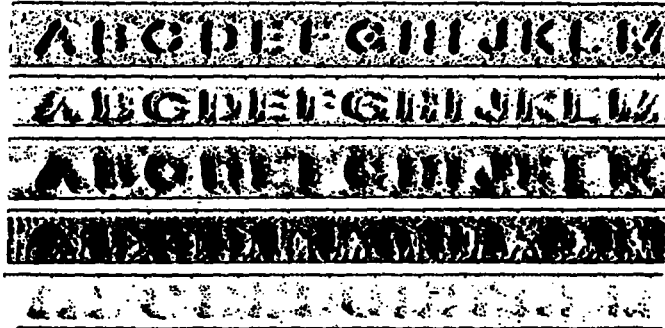


Figura 3. Los SEPs que se obtienen de los registros del área 3b.

Figura 4. Los SEPs que se obtienen de los registros del área 1 de la corteza.

Area 3b SA



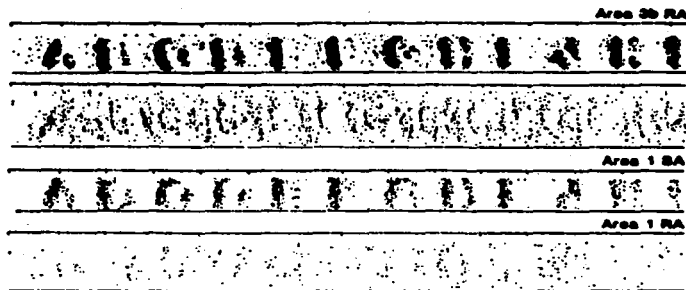


Figura 5. La traducción de algunos de los registros de cómo se producen los potenciales de acción cuando se forman los SEPs.

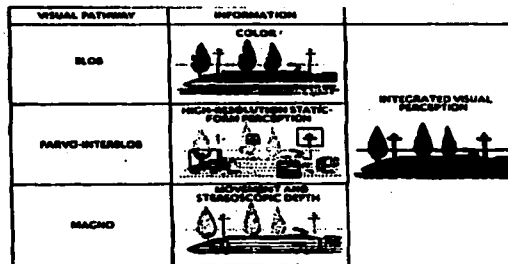


Figura 6. Livingston postula que el proceso de información visual está constituida por tres sistemas: la vía blob, que reconoce la forma del objeto, la vía parvo-interblob, que percibe el color y la vía magno, que percibe el movimiento y juzga la distancia a la que se localizan y se distribuyen los objetos.

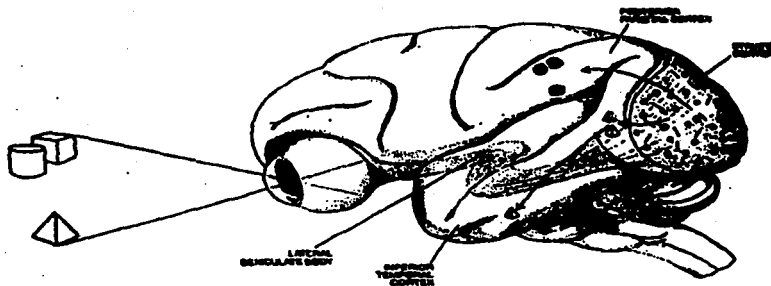


Figura 7. Imagen de cómo ilustran Mishkin y Appenzeller que la información visual es procesada de manera secuencial a lo largo de la retina, el cuerpo geniculado lateral, distintas áreas de la corteza visual y la corteza temporal inferior.

de tal manera que decir que toda mónada tiene percepciones significa que existen correspondencias entre la serie de sus conceptos y una serie de fenómenos. Es así como la noción típicamente epistemológica de percepción pasa a formar parte de la ontología leibniziana. En Leibniz la teoría del conocimiento forma parte integral de la ontología y se entiende a partir de ésta.

Bibliografía

- C *Opusculs et fragments inédits de Leibniz*, ed. por L. Couturat (Hildesheim: Georg Olms, 1961)
- E G. W. Leibniz. *Nuevos ensayos sobre el entendimiento humano*, ed. preparada por J. Echeverría Ezponda (Madrid: Editora Nacional, 1983).
- L G. W. Leibniz. *Philosophical Papers and Letters*, 2ª ed., ed. por Leroy E. Loemker (Dordrecht/Boston: Reidel, 1969).
- O G. W. Leibniz. *Escritos filosóficos*, ed. por Ezequiel de Olaso (Buenos Aires: Cáracas, 1982).

NEWTON: PERCEPCIÓN Y EXPLICACIÓN CIENTÍFICA

Nydia Lara Zavala*

Este trabajo aspira a esclarecer la posición de Newton frente a lo que puede y no puede decir la ciencia sobre las causas externas al sujeto y supuestamente generadoras de la percepción. Lo que desde un punto de vista histórico me interesa rescatar es la negativa que se desprende de diversos trabajos de Newton respecto a la tesis de que los denominados "objetos físicos" (átomos, fotones, partículas elementales, longitudes de onda, etc.) son las causas externas de nuestras percepciones sensibles. Así, el núcleo de mi interpretación de los trabajos de Newton consiste en mostrar que éste no concibe la relación entre los objetos físicos, nuestros órganos sensoriales y nuestra percepción cotidiana del mundo sobre las bases de una teoría causal de la percepción. Por otra parte, vale la pena señalar que Newton también sostiene que el hecho de que nosotros podamos estudiar y explicar científicamente cómo se afecta nuestra percepción cuando alteramos el funcionamiento de algún órgano sensorial o el cerebro mismo no nos autoriza a suponer que ellos son causa de dicha percepción. Para Newton, nuestros reportes sólo nos autorizan a afirmar que, si se altera o se daña un órgano sensorial o el cerebro, nuestra percepción sensorial de la realidad efectivamente quedará alterada o inclusive cancelada, pero el punto importante es que nuestros experimentos y observaciones, por sofisticados que sean, no nos revelan cuáles son las causas de nuestras percepciones. La idea es la siguiente: la condición de posibilidad para

*Laboratorio de Neurocomputación, Centro de Instrumentos, UNAM.

llevar a cabo cualquier tipo de experimento y explicación causal ya presupone que hay tal cosa como percepción. Dicho de otro modo, es sólo porque partimos de ella que estamos capacitados para decidir qué existe, qué es percibido y qué es lo que se altera o se modifica. Lo que quiere Newton decir con esto es que con las técnicas de la ciencia uno no encuentra qué es lo que hace que las cosas existan como las percibimos, sino que sólo se determina la manera como se relacionan y se afectan las cosas que sabemos que existen porque las percibimos.

La creencia en la existencia de un orden cósmico, susceptible de ser descrito matemáticamente por el intelecto humano está, sin duda, en el origen de ese modo de pensar que le da fundamento y forma a nuestra noción actual de ciencia. Ahora bien, el uso de las matemáticas como herramienta para derivar las características que definen a los objetos de la nueva filosofía experimental se constituye sobre la "nítida" distinción entre cualidades primarias y secundarias. Dicha distinción ganó su puesto en la historia gracias fundamentalmente a la influencia de Galileo, Descartes, Boyle y Locke. Originalmente fue pensada como una distinción de carácter ontológico y tuvo como una de sus primeras consecuencias no sólo la de desproveer de realidad a las cualidades secundarias, sino también la de hacer que se les concibiera como el producto de la actividad cerebral y psicológica del sujeto. Esto obligó a suponer que parte del quehacer del científico consistía en tratar de explicar lo que provocaban en nosotros la percepción de los objetos del mundo cotidiano.

Desde esta perspectiva, la distinción entre cualidades primarias y secundarias quedó enmarcada en el contexto de lo que se denomina la "teoría causal de la percepción", a la que podemos caracterizar como aquella teoría que sostiene que, de una u otra manera, los objetos físicos no sólo son objetos externos, distintos e independientes de la percepción, sino que son concebidos como la causa externa de nuestros datos sensoriales o representaciones

internas. Lo que tenemos son las "ideas" de los objetos inmediatos de nuestro conocimiento sensible o percibido. Desde este punto de vista, la percepción no es nunca de los objetos de la realidad física. Así, pues, la teoría causal de la percepción induce a creer que el mundo que percibimos no es el mundo real, sino sólo su apariencia sensible. La idea es que lo que percibimos cotidianamente es sólo el efecto que el mundo físico real genera en o gracias a nuestros sentidos, pero lo que aparece ante nuestros sentidos no es la realidad misma o en sí. Se trata más bien de las imágenes o representaciones internas que se producen en nosotros por la acción que provoca la realidad externa sobre nuestros sentidos.

Consideremos, por ejemplo, el caso de la visión. Actualmente se sostiene que ésta es causada (es decir, es el último eslabón de una cadena causal) por un complejo sistema de rayos u ondas luminosas, según sea el caso, que son emitidos o reflejados por la superficie de los objetos y que "viajan" del objeto al ojo del percipiente. Se supone que cuando un haz de luz llega a la retina se producen cambios químicos en los receptores, mismos que provocan que una serie de impulsos eléctricos viajen a lo largo de las fibras nerviosas hasta llegar hacia una de las dos áreas visuales del cerebro que reciben los impulsos nerviosos de la retina. Estos impulsos generan a su vez una actividad en esas áreas del cerebro así como en otras asociadas a ellas para provocar o causar en el percipiente la visión de un objeto.

Nosotros, si somos coherentes con lo que afirma esta teoría, habremos de decir que en principio no vemos los rayos u ondas emitidos o reflejados por los objetos. De hecho tampoco vemos los objetos que emiten o reflejan la luz (e.g., no vemos nuestro ojo por medio del cual vemos). Lo que vemos son simplemente

¹ Hirst, "Realism", The Encyclopedia of Philosophy, Vol. 7, Collier Macmillan Publishers, London, 1967.

los efectos que la emisión o reflejo de la luz provoca en nuestra retina. Se asume que el cerebro de alguna manera traduce dichos efectos como un conjunto de manchas coloreadas de determinada forma y tamaño, ubicadas en nuestro espacio visual como imágenes sensibles y que funcionan para nosotros como representaciones internas de los objetos externos. Pero hay que hacer notar que la distinción entre mundo externo y representaciones internas acarrea consigo la idea de que el mundo que cada uno de nosotros percibe tiene un carácter subjetivo y privado. Esto se fundamenta en la idea de que si, en efecto, la percepción es el producto de la actividad que literalmente se produce al interior del sujeto (por medio de su actividad cerebral y psicológica), lo que el sujeto percibe cotidianamente sólo puede ser percibido por el sujeto mismo. Hay, pues, desde el punto de vista de esta teoría, una duplicidad de mundos en cada sujeto: el que cada uno de nosotros percibe cotidianamente a través de sus sentidos y el que no percibimos directamente pero que suponemos causa en cada uno de nosotros nuestras percepciones sensibles.

La ciencia, como dijimos, se inicia dividiendo el reino de los objetos de la percepción sensible en dos mundos: el de las partículas equipadas sólo con propiedades matemáticas que se unen, separan o mueven según lo determinan las leyes de la mecánica y el de los colores, sabores, olores y demás cualidades que, por definición, no pertenecen al mundo físico. Esto, por toda una serie de supuestos condujo a la identificación de las partículas matemáticas con el mundo real y a la de las cualidades secundarias con elementos cuya existencia dependía por completo de la percepción del sujeto. Bajo esta perspectiva, la "percepción de los objetos externos" se concibe como la percepción de las sensaciones causadas por ellos. La teoría denominada "representacionista", defendida por gente como Galileo, Descartes y Locke, sin embargo, supone que aunque todo lo que percibimos se limita a la posesión de sensaciones, algunas de esas sensaciones

reflejan o reproducen de algún modo las propiedades físicas espacio-temporales de los objetos externos, en tanto que otras no.² El postulado de las primeras cumple con el propósito de garantizar que cuando menos algunas de nuestras descripciones de los objetos físicos coincidan con las características que realmente poseen los objetos externos. Esta posición queda ilustrada en las palabras de Galileo cuando éste afirma:

...yo juzgo que, si los oídos, las lenguas y las narices se eliminaran, la figura, los números y el movimiento sin duda permanecerían, pero no el olor o el sabor o el sonido, que, sin el animal viviente, no puedo creer que sean otra cosa más que nombres, así como las cosquillas no son otra cosa más que un nombre cuando la axila o la membrana de la nariz se eliminan.³

La separación que establecen los representacionistas entre lo que permanece y lo que se elimina al suprimir los órganos de los sentidos del "animal viviente" da lugar a que se piense que entonces el olor, el color o el sonido al igual que sensaciones como las cosquillas, el dolor o el placer no son reales. Pero al expulsar a las cualidades secundarias del mundo real se tiende a dotar a las partículas matemáticas de existencia propia. La realidad, desde esta perspectiva, no sólo se considera distinta e independiente del mundo que percibimos, sino que se piensa como la causa que produce en nosotros el cúmulo de sensaciones que cesarían de existir si el sistema que las produce se extirpa o se extingue. Ahora bien, Newton no sólo cuestionó la idea de que fuera necesario dar cuenta de la presencia de las cualidades secundarias en el mundo de la percepción cotidiana, apelando a la actividad que los objetos físicos provocan en nuestros sentidos

² Hirst, *Realism*, p. 81.

³ Galileo, *Opere*, IV, 336, ff.

sino que rechazó enfáticamente la tesis que pretende identificar a la realidad con los supuestos objetos externos de los que se habla en la física.

Para Newton, posiciones como la de Galileo no sólo no explicaban nada, sino que enredaban los temas de tal manera que se terminaba por no saber ni lo que la física afirma conocer ni cómo puede accederse empíricamente a la realidad. Desde este ángulo, Newton sostiene que el hecho incontrovertible es que:

En los cuerpos vemos sólo sus figuras y colores, oímos sólo sus sonidos, tocamos sólo su superficie externa, olemos sólo sus olores y saboreamos sus sabores.⁴

Que estos cuerpos son además extensos, duros, impenetrables, móviles e inertes es algo que Newton nos asegura que nosotros podemos reconocer sólo por medio de los sentidos y la experiencia. En su Regla III, por ejemplo, concretamente sostiene:

No tenemos otra manera de conocer la extensión de los cuerpos más que por los sentidos (...) Que muchos cuerpos son duros lo sabemos por la experiencia (...) Que todos los cuerpos son impenetrables no lo averiguamos por la razón, sino por la sensación (...) Que todos los cuerpos son móviles y que poseen ciertos poderes (que llamamos inercia) (...) lo inferimos sólo porque propiedades semejantes son observadas en los cuerpos que vemos.⁵

Partiendo de esto, Newton considera que si eliminamos los sentidos del "animal viviente", como propone Galileo, sencillamente no llegamos a ningún lado, ya que según él el supuesto

⁴ Newton, *Principia*, Book III, General Scholium, trad. de Motte, Univ. of California Press, USA, 1966, p. 456.

⁵ *Op. cit.*, Rules of Reasoning in Philosophy, Rule III, p. 398.

mundo externo que causa nuestras sensaciones no puede interpretarse más que como un supuesto metafísico redundante y que no tiene cabida en la ciencia experimental. En sus palabras:

Una necesidad metafísica ciega, que es ciertamente siempre la misma y en todos lados, no puede producir cambio alguno en las cosas.

Nosotros, según Newton, sabemos cómo son los objetos que existen porque éstos pueden presentarse a nuestros sentidos y los identificamos como objetos reales precisamente por sus características fenoménicas,

pero su sustancia interna no puede ser conocida ni por medio de los sentidos, ni por ningún acto reflexivo de nuestra mente.⁶

Querer, pues, determinar cómo serían los objetos del mundo si no los percibiéramos, implicaría que de alguna manera somos capaces de acceder al conocimiento de su "sustancia" y esto, en opinión de Newton, es un esfuerzo vano. Lo que conocemos no es nunca la sustancia de los cuerpos que percibimos.⁷ Conocemos únicamente la forma como se nos manifiestan a los sentidos. Decimos, por ejemplo, que un cuerpo puede definirse como algo que es duro, rígido, impenetrable e inerte. Si se nos pregunta cómo lo sabemos, contestamos que así es como se manifiesta a nuestros sentidos. Sin embargo, si se nos pregunta por qué se manifiesta así, toda nuestra sapiencia no alcanza para proporcio-

⁶ Newton, *Principia*, Book III, General Scholium, trad. de Motte, Univ. of California Press, USA, 1966, p. 546.

⁷ Newton, *Principia*, Book III, General Scholium, trad. de Motte, Univ. of California Press, USA, 1966, p. 546.

nar las razones de por qué percibimos a los cuerpos de determinada forma, con determinadas propiedades y con determinados colores, sonidos, olores, sabores, etc. Si estas propiedades o cualidades forman parte o no de la sustancia de los cuerpos es algo que nosotros no podemos determinar ni por medio de la reflexión ni por medio de la sensación. Percibimos las cosas que existen por medio de nuestros sentidos, pero al no poder determinar la sustancia de los objetos que percibimos no podemos determinar tampoco si lo que percibimos como sus propiedades o cualidades es el resultado de la actividad de una sustancia mental o material o cualquier otra cosa o combinación de cosas que se nos pueda ocurrir. Suponer, pues, que para percibir el mundo se requiere de la existencia de otro mundo pero sólo dotado de cualidades primarias no nos hace avanzar nada sencillamente porque no explica nada, lo cual, en palabras de Newton, "equivale a no decir nada". En cambio, continúa él diciéndonos:

...derivados o tres principios generales del movimiento a partir de los fenómenos para decir a continuación cómo se siguen de esos principios manifiestos las propiedades y acciones de todas las cosas corpóreas, habrán de constituir un gran paso en la filosofía, aunque las causas de esos principios aún no se hubiesen descubierto.⁶

Lo que yo deseo sostener es que, aunque Newton en cierto sentido acepta la necesidad de distinguir entre cualidades primarias y secundarias, su admisión no lo compromete a asumir que se trata de una distinción ontológica que deja fuera de la realidad el mundo constituido por nuestros objetos de percepción. Para él, diferenciar entre uno y otro tipo de cualidades parece que responde más bien a una estrategia de corte epistemológico y práctico, sin otro propósito que el de apresar y definir mejor las

⁶ Newton, *Optica*, p. 346

características universales que consideramos necesarias para lograr identificar a los objetos corpóreos con que nos topamos en la vida real.⁶ Pero, si Newton es coherente, todo parece indicar que para él, el hecho de que el color, el sabor, el sonido, el olor o la textura de un cuerpo no se incluya en la definición de 'cuerpo físico' se debe simplemente a que este tipo de cualidades no se requiera para dar cuenta del fenómeno de ciertos cambios de los cuerpos y que Newton juzga como aquello que realmente interesa a la física estudiar. En sus *Principia*, por ejemplo, una y otra vez sostiene que el rango de acción de la física se limita a dar cuenta del fenómeno del "movimiento" y para explicar el movimiento de los cuerpos es evidente que no necesitamos introducir todas las propiedades manifiestas fenoménicamente en los cuerpos. En su *Óptica*, sin embargo, él introduce el color como característica esencial para poder explicar el comportamiento de la luz.

Si mi lectura de Newton es correcta, lo que podemos suponer es que él separa a las cualidades secundarias de las primarias simplemente porque lo que él pretende es lograr una definición de 'cuerpo físico' lo suficientemente precisa como para poder proporcionar un lenguaje y un mecanismo que nos haga posible distinguir claramente y sin ambigüedad cuándo estamos hablando de un cuerpo y cuándo no. Esto tiene vital importancia para la visión newtoniana de la física, ya que la misma definición de cuerpo físico le va a permitir plantear que hay cosas que existen en el mundo que percibimos y que estudia la física que por definición no son "corpóreas". El espacio vacío (o lo que él caracterizó como "éter") y la luz son, para Newton, fenómenos de este tipo. Lo que en todo caso quiero dejar en claro es que, desde la perspectiva newtoniana, si se define 'cuerpo' como algo

⁶ McMullin, *Newton on Matter and Activity*, Univ. of Notre Dame Press, USA, 1978, p. 9.

duro, rígido, impenetrable e inerte, lo que se quiere establecer es que cualquier cosa que no presente esas propiedades tendrá que ser calificada como no corpórea, mas no por eso como algo no real, puesto que identificamos lo real con el fenómeno observado. Asimismo, la identificación de un cuerpo por medio de sus cualidades primarias no significa que el cuerpo identificado no posea las demás cualidades que les son propias a los objetos corpóreos que percibimos cotidianamente y que sin duda son los mismos cuerpos que la física pretende estudiar.

Por ahora lo que me interesa señalar es que, aunque Newton nunca publica su formulación de la definición del término 'fenómeno', es claro que en toda su obra él concibe al fenómeno como el único objeto de estudio de la ciencia experimental. Si esto es cierto, entonces la propuesta de que existe un mundo externo y distinto del que percibimos, cuya actividad causa en nosotros percepciones, es para él el tipo de hipótesis que no tienen cabida en la ciencia experimental. En palabras suyas:

...cualquier cosa que no sea deducida desde el fenómeno debe ser llamada una hipótesis; y las hipótesis, sean metafísicas o físicas, referirán a cualidades ocultas o mecánicas, no tienen lugar en la filosofía experimental.¹⁰

Bien entendido lo que esto quiere decir es que los cuerpos físicos de Newton ni pueden ser diferentes física o metafísicamente de los fenómenos (que sí percibimos) ni pueden estar dotados de características o poderes ocultos gracias a los cuales los percibimos. De hecho, el cuerpo para Newton se concibe sólo como el elemento que podemos identificar objetivamente por medio de los sentidos para detectar, a través de la observación

¹⁰ Newton, *Principia*, Book III, General Scholium, trad. de Motte, Univ. of California Press, USA, 1966, p. 547.

de sus movimientos y cambios, cuándo hay una causa presente que los altere o modifique. Dicha causa, empero, no sólo no es externa al fenómeno, sino que nosotros sabemos que existe como fenómeno en la medida en que somos capaces de observar los cambios y alteraciones que sufren los cuerpos cuando esas causas están presentes. Si la causa es física, entonces debe tener propiedades o manifestaciones fenoménicas. En Newton, pues, no sólo no encontramos una distinción entre el mundo físico y el mundo de la percepción cotidiana, sino que de hecho encontramos una identificación entre el mundo que estudia la física y el de la percepción cotidiana. De esto, lo que puede inferirse, yendo en contra de toda una tradición, es que lo que Newton está tratando de hacer es enmarcar o circunscribir el terreno de la física en el estudio de los fenómenos. Lo que Newton piensa y defiende es que nosotros podemos (y tenemos que) identificar a los cuerpos sólo por sus manifestaciones fenoménicas, al igual que podemos identificar las causas o principios generales que mueven a esos cuerpos desde los fenómenos. Y aunque esas causas, para Newton no son ni pueden ser catalogadas como 'corpóreas', de ninguna manera significa que no forman parte o que están fuera del fenómeno. No hay, pues, la necesidad de postular la existencia de algo aparte del fenómeno para explicar el fenómeno ya que, como bien lo especifica Newton, la postulación de entidades otras que fenómenos para explicar la existencia de los fenómenos sencillamente no forman parte de la física. Si esto es correcto, entonces tenemos que aceptar que, cuando menos para Newton, la teoría causal de la percepción tal como se ha formulado no es una teoría científica sino una hipótesis "metafísica".

Aunque, como se dijo, Newton nunca publicó una definición del término 'fenómeno', en sus notas y diferentes esbozos encontramos que para él dicho término tiene un significado muy amplio, ya que abarca todas y cada una de las cosas que percibimos. El nos dice, por ejemplo:

Voy a llamar fenómeno a cualquier cosa que pueda ser vista y ser percibida sea cual sea el objeto percibido, refiera a las cosas externas que nos son conocidas por medio de los cinco sentidos, o a las cosas internas que contemplamos en nuestras mentes por el pensamiento. Como el fuego es caliente y el agua es húmeda, y el oro es pesado, y el sol es luz, yo soy y yo pienso. Todo esto son cosas sensibles y pueden ser denominadas fenómenos en un sentido amplio.¹¹

En el mismo tono, Newton nos dice en otro pasaje:

...yo no tomo por fenómeno sólo a las cosas que se nos hacen conocidas por los cinco sentidos externos, sino también aquellas que contemplamos en nuestras mentes cuando pensamos: como son, yo soy, yo creo, yo entiendo, yo recuerdo, yo pienso, yo deseo, yo no deseo, yo estoy sediento, hambriento, contento, triste, etc. Y aquellas cosas que no se siguen de los fenómenos ni por demostración ni por el argumento de inducción, lo tomo por hipótesis.¹²

Pero Newton, me parece, además de las cosas que pueden ser vistas y percibidas por los sentidos internos y externos, incluye dentro de la categoría de "fenómeno" no sólo cuerpos y sensaciones, sino hechos.¹³ Los seis Fenómenos que siguen a las Reglas de sus *Principia*, por ejemplo, nos hablan no sólo del sol, los planetas y sus satélites. Newton también considera como fenó-

¹¹ Yo tomé esta definición de Peter Achinstein, quien refiere a la traducción que hace McGuire, "Body and Void in Newton's *De Mundi Systemate: Some New Sources*", *Archive for History of Exact Sciences* 3, 1966, p.230-239.

¹² Cohen, I. Bernard, *Introduction to Newton's 'Principia'*, Harvard Univ. Press, USA, 1978, p. 30. Según Achinstein, además de la cita anterior y la que proporciona Cohen existen dos definiciones adicionales entre los manuscritos de Newton que él no cita porque sostienen básicamente lo mismo. Yo no he tenido acceso a ellas.

¹³ ver Achinstein, *Newton's Corpuscular Query*.

menos a las órbitas de los planetas y sus satélites, a las áreas que estos cubren, a los períodos de tiempo que tardan en recorrer sus órbitas, etc. Órbitas, áreas, períodos y demás cantidades quedan, pues, agregados a la lista de la referencia del término 'fenómeno', ya que su existencia es o puede ser establecida por la observación y la experiencia o la experimentación.¹⁴ Pero si todo lo que puede ser establecido por la observación y la experiencia forma parte de la categoría de fenómeno, uno esperaría que la inercia, la gravedad, los rayos de luz, los átomos o las partículas elementales que ahora sabemos que los componen, también pudiesen ser considerados por Newton como fenómenos. Sin embargo, no es este el caso. La inercia, la gravedad, las partículas o los rayos que él sostiene que componen a la luz son, para Newton, el resultado de inferencias y, aunque ciertamente pueden derivarse de los fenómenos que percibimos, sólo tienen cabida como categorías o entidades teóricas. Newton, si no me equivoco, distingue entre fenómeno e inferencia con el propósito específico de explicar el comportamiento, la modificación o los cambios que podemos observar en los fenómenos sin que lo que se infiere del fenómeno le acarree algún tipo de compromiso ontológico. Yo creo que esto es lo que le permite hablar, por ejemplo, de la gravedad en el caso de los cuerpos o de los rayos de luz en el caso de los colores, como simples "entidades teóricas", cuya "existencia" puede ser postulada sólo porque y en la medida en que contribuyen a explicar la manera como vemos que se comportan los cuerpos, la luz, etc.

El movimiento de un planeta, por ejemplo, es algo que podemos observar como fenómeno. La descripción de ese fenómeno en relación con las fuerzas que suponemos actúan para moverlo nos explica el fenómeno, pero no garantiza que sólo su concepción de la fuerza de gravedad corresponda a la única manera posible

¹⁴ Achinstein, p.137

de dar cuenta de ese fenómeno. Igual sucede con la noción corpuscular de los rayos de luz. A él le pareció que lo más coherente era explicar el comportamiento de los rayos de luz pensando en ellos como corpúsculos (no corpóreos) que se desviaban al pasar de un cuerpo o medio transparente a otro o se reflejaban cuando chocaban con la superficie de un cuerpo. Pero Newton nunca desechó la posibilidad de que la refracción y la reflexión de los rayos de luz también pudiesen ser explicados concibiendo a los rayos de luz como ondas.

Lo que en todo caso hace Newton consiste en establecer ciertas relaciones entre, por ejemplo, la fuerza de gravedad y los movimientos planetarios, pero lo único que él pretende estar haciendo es dar cuenta de los movimientos planetarios a partir del postulado de la fuerza de gravedad. Por medio de la fuerza de gravedad sin duda se explica el movimiento planetario y se puede decir con mucha precisión cómo actúa la fuerza de gravedad sobre los cuerpos. Del fenómeno del movimiento Newton pensó que era lícito inferir que existen ciertas fuerzas que provocan esos movimientos, pero cuando se le presionó para que contestara qué era eso que causaba esas fuerzas o qué o cómo se producían, Newton una y otra vez respondió que no sólo no lo sabía, sino que él ni proponía ni construía hipótesis a ese respecto. Para él, la causa del movimiento planetario podía muy bien explicarse con la noción de "fuerza", pero la causa, la localización, el origen, etc., de esa fuerza era algo que él sabía que no podía explicar científicamente. Esto, sin embargo, no le impidió elucidar sobre qué podía ser aquello que provocaba la gravedad. Empero, siempre consideró que cualquier cosa que él pudiese decir al respecto debía ser entendida como una simple "hipótesis", en su sentido peyorativo (es decir, metafísico) del término. Lo que me interesa dejar claro de todo esto es que Newton, aunque reconoce que no sabe qué causa la manifestación de las fuerzas gravitatorias o qué es la luz y qué la produce, ninguna causa ni ninguna cosa que se

infera del fenómeno puede concebirse como algo que exista con independencia del fenómeno que se observa. En este espíritu, Newton también propone, sobre la base de nuestras observaciones de su comportamiento, que se infera, e.g., que la luz consta de partes. En sus palabras nos dice que las partes de la luz se inferen de su comportamiento

...porque en el mismo lugar puedes parar aquella que viene en un momento y dejar pasar aquella que viene un poco después; y al mismo tiempo puedes detenerla en cualquier otro lado y dejarla pasar en cualquier otro. De aquí que la parte de luz que se detiene no puede ser la misma que aquella que se deja pasar. La parte más pequeña de luz o la parte de la luz, que puede ser detenida sólo sin el resto de la luz, o propagarse sola, o hacer o sufrir cualquier cosa sola, sin que el resto de la luz lo haga o lo sufra, yo lo llamo rayo de luz.¹⁵

Pero, como se puede claramente apreciar, Newton no propone que la luz sea corpuscular, sino sólo que consta de partes, a saber, los rayos de luz. Ahora bien, el hecho de que él proponga la existencia de rayos luminosos como las partes constitutivas de la luz sólo tiene como función explicar una gran variedad de manifestaciones fenoménicas de la luz. Sus experimentos sobre colores, por ejemplo, lo que explican es la relación entre la percepción de determinado color y el grado de refrangibilidad o desviación de las partes o rayos de luz. Pero Newton no está pensando que los rayos puedan ser entendidos como aquello que causa nuestra percepción del color, sino sólo que existe una relación entre los colores y la manera como se observa cómo se separan las distintas partes o los distintos rayos que está suponiendo que componen a la luz. Si esto se entiende, creo que es

¹⁵ Newton, *Optics*, Definition I, William Benton, Publisher, The Great Books, Encyclopaedia Britannica, Inc., 1952, p.379.

fácil apreciar que para Newton, tanto la luz como los colores que observamos cuando se descompone la luz al pasarla a través de un prisma son fenómenos, en tanto que los rayos luminosos son propuestos como entidades teóricas que sólo tienen la función de explicar el comportamiento que podemos observar en esos fenómenos. En otras palabras, para Newton las entidades teóricas, como son los rayos luminosos, las partículas elementales, o las fuerzas, no son ni pueden ser entendidos como la explicación causal de la existencia de los fenómenos que percibimos y esto por una razón muy simple: ellos sólo son aquellos elementos cuya existencia se postula con el único propósito de ofrecer una explicación coherente e inteligible del comportamiento que observamos en los fenómenos. Esto quiere decir que la explicación científica de un fenómeno ni es ni equivale a la explicación de la causa que lo produce. Si esto es cierto, la moraleja que podemos extraer es de suma importancia, a saber, que lo que la ciencia explica no rebasa nunca los límites fijados por el reino de la percepción.

VER LA LUZ: LOS RAYOS DE NEWTON, LA MEMORIA DE GOETHE Y LA PERCEPCIÓN DE LA CIENCIA

Dennis L. Sepper*

Al principio de 1992 se anunció que un experimento en proceso, a gran profundidad bajo tierra en Gran Sasso, en las montañas de la región de los Abruzzo en Italia, había confirmado la existencia de neutrinos de baja energía, partículas subatómicas cuya existencia se predijo en la década de 1930 para dar cuenta de la distribución de energía producida por la fusión de los átomos de hidrógeno en el sol. Experimentos similares se han llevado a cabo en otros lugares (p. ej., en una vieja mina de sal bajo el lago Erie en los Estados Unidos), pero el emplazamiento del Gran Sasso tiene la ventaja de estar mucho mejor aislado de otros fenómenos cósmicos, de tal manera que los neutrinos son las únicas partículas subatómicas que es probable que penetren ahí.

Los neutrinos se cuentan entre las partículas más elusivas; tienen cero masa de reposo y aun cuando, conforme a la teoría, miles de millones de ellos están pasando a través de cada centímetro cúbico de la tierra, por segundo, sólo interactuará con la materia una fracción infinitesimal de ellos. El experimento del Gran Sasso está ideado para capturar una fracción muy pequeña de neutrinos de baja energía —realmente, la parte alta de la porción baja del espectro de energía— en una enorme tina de galio líquido. Cada captura tiene como resultado la conversión, en gas, de uno de los átomos de galio que, tras muchos meses de mantener el experimento, puede evacuarse por la parte alta de la

*University of Dallas.

PERCEPTION AND SCIENTIFIC EXPLANATION

NYDIA LARA

Laboratorio de Neurocomputación
Centro de Instrumentos
Universidad Nacional Autónoma de México
Apartado Postal 70-186
México, 04510, D.F.
México

ABSTRACT

The aim of this work is to elucidate that the idea that through the study of our sense organs and our brain it is possible to give a satisfactory answer of the causal source of our sense perception, not only is a mistaken idea, but the origin of all kinds of philosophical puzzles. The point I want to stress in this paper is that even when we are able to study and explain, scientifically, how perception is affected when some senses organs or the brain is altered in its functions, it does not allow us to support that we perceive the world through them.

Our scientific reports only authorize us to assert that, if our sense organs or the brain are altered or damaged, our perception of the world in fact is going to be altered or even canceled, but the interesting point is that our experiments and observations, however sophisticated they might be, do not make us known the causal source of our perceptions. The main idea is the following one: the possibility to accomplish any kind of experiment and causal explanation already presupposes that there is such thing as perception. In other words, it is only because we perceive that we are able to decide what exists, what is perceived, and what is that which is altered or modified. With this I want to say that through scientific tectonics it is not possible to find what makes things to exist the way we perceive them. They only are able to determine the way things are related or affected that we already know that exist because we are able to perceive them. Nevertheless, the so called 'causal theory of perception', upheld today by many neurophysiological studies, induce us to believe that the world we perceive is the effect that the external world produces in our sense organs and our brain.

KEYWORDS: Perception, causal theory of perception, causal source of perception, causal explanation.

INTRODUCTION

Nowadays, not only philosophers and psychologists, but even scientists and engineers have been concerned with some conceptual puzzles in regard to the nature (e.g., mental or physical) of the processes involved in what has been called the 'study of perception'. These conceptual puzzles arise, in part, because there is not a clear understanding of the term

'perception', and, in part, because people tend to adopt, as a correct scheme to grasp this elusive concept, the metaphysical doctrine known as the 'causal theory of perception'. This theory can be characterized as the doctrine that sustains that the objects that conform the world not only are external, different and independent of our sense perception, but are to be considered as the causal source of our sense data or internal representations.

According with this theory, sight, for example, has to be considered as the final stage or effect caused by a complex system of wavelengths that are emitted or reflected from the object's surface to the percipient's eyes. Combining well grounded physical terms with metaphysical assumptions the theory asserts that, when different wavelengths strike the eyes, the light causes chemical changes in the receptors, which, in turn, cause electrical impulses to pass along the nerve fibers which lead from the receptors to one of the two visual receiving areas of the brain. The impulses set up activity there and in certain other association areas of the brain to provoke or cause in the percipient the visual image of the object (Hirst, Perception and the External World, p. 10).

Now, many of us are more than familiar with this kind of explanations, and several people tend to believe that this causal chain has to be considered as the point of departure of any scientific explanation that wants to assert the sources of our perception. Nevertheless, if we look closer and try to be coherent with what the theory wants to assert, the final result is that there are elements that conform this causal chain that are in fact utterly unexplainable. One of them is the access to the external world described in physical terms, the other concerns the source of perception itself caused by brain activity. Let us try to follow why this is so.

According with the causal theory of perception, we have to admit that, in principle, we do not see the wavelengths emitted or reflected from the objects. In fact, neither we see the object that emits or reflects the light (*e.g.*, we do not see the eye through which we see). What we see, in conformity with the theory, is merely the effect that the emission or reflection of light produce in our retina. Furthermore, it is not explained, but merely assumed that somehow the brain is able to translate those effects as a set of colored spots endowed with certain shape and size that can be located in our visual space as sensible images that work for us as the internal representations of external objects.

Note that the theory tend to multiply everything. There is a physical description of the external world and its common sense description; there is a external space where the physical objects inhabit and the perceived space where our perceptual objects inhabit. This take us at least to two forms of dualism: one is the so called 'epistemic dualism' that has to confront the doubts concerning the possibility to have perceptual access to the external world; the other is the so called 'ontological dualism' or the 'mind-body dualism' that, among other things, has to confront the problem concerning how is it possible for a physical cause to produce in us a non physical effect.

Now, the mental aspect that surround the term under study in connection with the conceptual split that the causal theory of perception establish between the external world

and its internal representations not only are responsible for the problematic distinction between the physical and the mental in both, its epistemic and ontological versions, but of the astonishing idea that sustains that our perception of the world is something that takes place inside our heads. This idea, as I will try to show, not only is obscure, but imprecise, in the sense that the phrase "inside our heads" originally was coined to express the metaphysical statement that asserts that our perception of the world is something mental and, as such, the phrase wants to express something odd: the location of perception in our minds. This idea is erratic because it seems that both, mental states and minds accept spatial descriptions, which is utterly absurd. Nevertheless, if we take this nonsense as an empirical statement (which is common when we are not aware of the dangers that characterize metaphysical statements), it might suggest that, if we could be able to explore what is going on, verbatim, inside our heads, we might be able to "discover" where and how exactly perception takes place in our heads.

Now, we usually have no difficulties in finding the place of a thing perceived. The problem here is that the so called 'neurophysiological studies of perception' do not want to find the location of something perceived (e.g., a brain, its nerves, neurons, chemical elements, etc.), but of perception itself. This pretension is thoughtlessness when we understand that to perceive is a verb not a thing that accept to be perceived. Nevertheless, the attempt to locate perception somewhere inside our heads has been accepted as a coherent possibility and has been responsible to create the illusion that one of the task of science is to study how perception takes place in the brain. But let us try to understand how this gibberish start.

ABOUT THE TERM 'PERCEPTION'

The term 'perception' undoubtedly is related with verbs like 'seeing', 'touching', 'tasting', 'hearing', and 'smelling'. All these verbs are to be considered as behavioral verbs, and all of them express performances strongly linked with what is known to be an 'ostensive definition'. We usually define a term ostensibly when we combine a perceptual verb with some form (physical or verbal) of pointing to the object.

Now, perceptual verbs when combined with some form of pointing to objects usually express some kind of behavioral performance that has to be done with the object pointed (e.g., see the apple, touch the material, taste the food, hear the sound, smell the perfume, etc.). If required, we can corroborate that the performance is fulfilled employing phrases like, e.g., "I see it", "I feel it", "I taste it", "I hear it", "I smell it", etc. Furthermore, we can extend the results of our performance adding new descriptions concerning the characteristics or properties found in object, flavors, sounds or smells (e.g., this is red, this is soft, this is sweet, that is a melody, this is a floral smell, etc.). The term 'perception', hence, seems to be related not only with verbs (e.g., seeing), but with commands (e.g., see that), confirmations (e.g., I see it), and description (e.g., this is red). All these conform the complex meaning of the term 'perception' and, as it is easy to appreciate, the same term sometimes is employed to express commands, sometimes performances, sometime assertions, sometimes objects, and

even sometimes it is used to describe properties or characteristics found in objects, sounds or smells.

Despite the different uses and improvements of the term, we tend to think that when we employ the same word to express miscellaneous but related things, the term must have a common reference. Hence, when philosophers like Descartes and Locke try to capture what they considered to be the "essence" of the term 'perception' they simply couple ostensive descriptions of objects and first person reports and assumed that the combination of both was the expression of some kind of mental or internal state produced by the object perceived. From here the idea that our perception is our mental apprehension of the external world.

Nevertheless, note that this presumed mental state arises because it is assumed that when we see an object and describe it, let us say, as red, the description "this is red" appears to be equivalent to the assertion 'I see it red'. This peculiar assertion is enough to give rise to two interrelated misconceptions. One of them is that the phrase "I see it" (that commonly is understood as our verbal way to express the confirmation that the performance stated by the coupling phrase "see that" is fulfilled) under this interpretation can be taken to be the report of a "mental" or "internal" state produced in us by the object perceived. The other one is that terms like 'red' now seems to acquire an ambiguous status, in the sense that under this interpretation when we point to an object and describe it, e.g., as red, we do not know for sure if the term 'red' wants to express the way we subjectively perceive the color of the object, or if it is used to assert merely the color of the object. This ambiguity provokes doubts about the reference of the term 'red', in the sense that, if we want to put the reference of that term as the description of my private mental state, then, the term automatically loses its public reference and, accordingly, its public meaning. This is so because now there seems to be no way to find out if the 'red' I say to perceive in the object is the same perceived by someone else. From here the idea that whatever is known is relative to the mind that knows it.

Now, when people confront this puzzle they usually use the recourse to put back the reference of the term in the object that is presumed to be the source of our perception. This change of reference is used in order to sustain that the object that exist outside us has its own properties, and those properties are able to produce in each one of us similar effects. The problem here is that we already accept that we have not a direct access to the properties of external objects. That is, in principle, we already accept that we only have access to the effects that we suppose external objects produce in each one of us. The result of this assumption is that we not only cancel in advance the possibility to compare the effects that external objects produce in each one of us, but also the possibility to know the characteristics of external objects that generate in us our perceptions. This take us into the absurd situation where we are unable to corroborate causes with effects, and, if this is so, then we have to accept that we do not have the empirical mechanisms that authorize us to give any kind of scientific explanation about the source of our perception.

Nevertheless, and despite this last conclusion, some scientists tend to defend the validity of the causal chain asserted by the causal theory of perception in the grounds that if we brake any link of that chain, our perception of the external world is automatically interrupted. What they do not consider is that that causal chain is a myth that has not any chance of empirical support.

Now, it is undeniable that if our sense organs or the brain are altered or damaged, our perception of the world in fact is going to be altered or even canceled. Nevertheless, our experiments and observations, however sophisticated they might be, do not allow us to explain why, how and where we perceive. The main idea is the following one: the possibility to accomplish any kind of experiment already presupposes that there is such thing as perception. This means that it is only because we are able to perceive that we are able to decide what is that which is altered or modified when we report or observe that our perception is altered or canceled. Note, however, that the correlation in this case is made between the impossibility to perceive something (that is, the impossibility to respond correctly to some object, sound or smell) and the damage or alteration detected (through our experiments and direct observation) in our sense organs or brain.

An inattentive mind might think that the correlation I am asserting between altered perception and physical damage is the same asserted by the causal theory of perception. But this is not the case. I am asserting that the physical damage only is able to explain the alteration of perception, and not the source of our perception. Whereas the causal theory of perception is trying to find out how external object generate our internal or mental representations of things through our sense organs and our brain. In order to do that, the causal theory of perception has to assume that the external objects are the causal source of our perception and, hence, the theory has to find a correlation between our internal representations and the external objects, which, as we already see, is not only an impossible enterprise, but different that the scientific project.

Percepción, metafísica y explicación científica

Nydia Lara Zavala
Laboratorio de Neurocomputación
Centro de Instrumentos
UNAM

Es digno de observar que en la actualidad no sólo los filósofos y los psicólogos, sino los biólogos y hasta los ingenieros están interesados en discutir la problemática "naturaleza" de los procesos (físicos y/o mentales) que se supone generan nuestra percepción sensorial. Empero, y como trataré de hacer ver en este trabajo, cuestiones referentes a la naturaleza de algo, aunque a veces parecen cuestiones empíricas, son preguntas de carácter conceptual. Esto es, discusiones acerca de la naturaleza de la percepción tienen que ver más con la necesidad de definir claramente ese término que con cuestiones experimentales. Lo que quiero decir es que cuando alguien se pregunta por la naturaleza de la percepción, lo que requiere no es descubrir experimentalmente si la percepción es física o mental, sino que lo que requiere es definir las características que debe tener ese concepto para poder decidir si dichas características gramaticalmente refieren a algo físico o a algo mental (Hospers, An Introduction to Philosophical Analysis, p.39).

Ahora bien, parte de lo que yo quiero explorar en este trabajo es que el origen de múltiples discusiones recientes en torno a la naturaleza de la percepción responde al hecho de que la gente involucrada en su estudio asume, en primer lugar, que la percepción debe ser entendida como el resultado de un proceso y, en segundo lugar, que el esclarecimiento de la naturaleza de los procesos que se supone generan nuestras percepciones pueden ser definidos por la vía empírica. Sin embargo, y como lo trataré de hacer ver, esta idea sólo puede ser sostenida debido a ciertas confusiones conceptuales implicadas en el pensamiento pre-newtoniano (o pre-científico). Dichas confusiones se gestan al tratar de definir las propiedades físicas de los objetos apelando a la entonces nueva y a primera vista "nitida" distinción entre cualidades primarias y secundarias. Esta distinción ganó su puesto en la historia gracias a la influencia de Galileo, Descartes, Boyle y Locke. Estos pensadores, al definir 'cuerpo físico' sólo en términos de sus cualidades primarias, confundieron una distinción conceptual con una ontológica. Esta confusión provocó que se pensara que al definir un concepto nuevo se estaba haciendo referencia a una clase "especial" de objetos recién descubiertos. Esto los llevó a cometer tres errores, a saber:

- 1) asumir que si al construir el concepto o al dar la definición de 'cuerpo físico' se requería únicamente apelar a propiedades primarias, es decir, no se requería incluir en la definición a todas las características observables en los objetos cotidianos, ello se debía a que los cuerpos físicos, a diferencia de los objetos cotidianos, no poseían cualidades secundarias.
- 2) asumir que el mundo que describe la física sólo en términos de cualidades primarias debía ser comprendido como un mundo real.

3) sostener que si ese mundo real no poseía cualidades secundarias, entonces la realidad que describe la física evidentemente tenía que ser concebida como una realidad distinta e independiente de la percibida, en vista de que la realidad física y la percibida simplemente no coincidían.

El resultado de estos tres errores fue provocar una distinción espúrea entre la realidad que define y estudia la física y la realidad cotidianamente percibida. Distinción que ha sido usada para convencernos de que la percepción cotidiana de la realidad es algo así como una "aprehensión mental" que tienen los percipientes de los objetos físicos. Dicha idea no sólo o hace que la percepción se conciba como una especie de actividad mental, sino que dicha actividad se interprete como el resultado de la interacción de dos mundos: el que describe la física, básicamente en términos de cualidades primarias, y el cotidiano que incluye los colores, sabores, olores y demás cualidades que, por definición, no pertenecen al mundo físico.

De todo esto fue surgiendo la así llamada 'teoría causal de la percepción', a la que podemos caracterizar como aquella teoría que, montando un error sobre otro, trata de relacionar el mundo físico con el mental asumiendo, en primer lugar, que lo que percibimos normalmente no son los objetos físicos, sino que más bien tenemos representaciones mentales de ellos y, en segundo lugar, que esas representaciones mentales deben ser entendidas como si fuesen las ideas causadas o generadas por la acción que producen los objetos físicos en nuestros órganos de los sentidos y el cerebro. Desde esta perspectiva el uso del término 'percibir' se convierte no sólo en algo extraño sino problemático, ya que la percepción queda caracterizada, en esta intrincada construcción metafísica, como una representación mental de objetos físicos al tiempo que se asume que dicha representación mental debe ser entendida como algo ocasionado o causado por una serie de procesos físicos. Esto último es problemático porque, así expuesto, parece entonces que son los procesos físicos los que de alguna manera deben ser entendidos como los responsables de producir o causar en los percipientes representaciones mentales. Esto, empero, es sumamente extraño porque normalmente si se dice de algo que es mental es justamente para señalar que no es físico y ahora resulta que es lo físico lo que produce lo mental. O sea, lo físico es susceptible de producir algo no físico en el percipiente, a saber: representaciones mentales, siendo dichas representaciones además diferentes cualitativamente a los objetos físicos: que entran en el proceso causal de la percepción. El problema aquí es que una vez que se acepta que lo que percibimos no son objetos físicos resulta difícil saber no sólo qué son los objetos físicos mismos, sino también qué relación guardan los objetos físicos con las percepciones que se supone que ellos producen.

En efecto, si tomamos en cuenta toda clase de representaciones mentales, como por ejemplo ilusiones, sueños y alucinaciones, se vuelve no sólo imposible saber si efectivamente existen o no objetos físicos cuando los percibimos, sino que aún suponiendo que están allí causando nuestras percepciones, no parece que sea posible determinar con sólo los datos percibidos, qué características efectivamente pertenecen a los objetos físicos y cuáles son

sólo un producto secundario de nuestras imágenes mentales. Tratando de solucionar este problema, Galileo, por ejemplo, sostiene:

...yo juzgo que, si los oídos, las lenguas y las narices se eliminaran, la figura, los números y el movimiento sin duda permanecerían, pero no el olor o el sabor o el sonido, que, sin el animal viviente, no puedo creer que sean otra cosa más que nombres, así como las coquillas no son otra cosa más que un nombre cuando la axila o la membrana de la nariz se eliminan (Galileo, Opere, IV, 336, ff).

Ahora bien, la separación que establece Galileo entre lo que permanece y lo que se elimina al suprimir los órganos de los sentidos del "animal viviente" hace que se interprete eso que permanece y que se identifica con los objetos físicos, primero, como algo externo, distinto e independiente de lo que percibimos y, segundo, como aquello que causa en nosotros, a través de nuestros sentidos, el cúmulo de sensaciones que cesarían de existir si el sistema que las produce se dañara o alterara. De aquí la idea de que nuestras representaciones mentales son el producto de nuestras sensaciones.

Por otra parte, parece estar implicado que lo que se podría considerar como la "percepción de los objetos externos" ahora pasa simplemente a ser la percepción de las sensaciones causadas por ellos. Desde esta perspectiva, naturalmente, se concluye que lo que normalmente percibimos no es nunca la realidad en sí misma y que lo que tenemos es más bien las imágenes o representaciones "alteradas" o "deformadas" que se supone se producen en nosotros por la acción que provoca la realidad externa en nuestros sentidos. Esto último de alguna manera genera la idea de que tanto la percepción normal como la anormal forman parte de una única línea de investigación y que su estudio se conciba como algo inevitablemente ligado al estudio de nuestros órganos sensoriales y, por extensión, al cerebro.

Desafortunadamente, esto de alguna manera no sólo ha dado pié, sino que ha permitido que dos cuestiones distintas constantemente se confundan. Las cuestiones a las que me refiero son, por una parte, una de carácter genuinamente científico y una de naturaleza metafísica. En relación con la primera se trata de averiguar cómo se modifica la conducta visual, auditiva, olfativa, gustativa o táctil cuando se alteran, dañan o extirpan nuestros órganos de los sentidos o partes del cerebro; en cuanto a la segunda, lo que se trata de hacer ver es que eso que los científicos descubren que es capaz de alterar, dañar o cancelar la conducta es a su vez responsable de causar o evitar que se produzcan en nosotros las imágenes o representaciones mentales del mundo externo. Sin duda alguna, en ambos casos lo que se investiga tiene que ver con la percepción, pero mientras que para el científico 'percibir' significa poder ver, tocar, oír, gustar u oler los objetos, sonidos, sabores u olores que están en nuestro entorno, para el metafísico tanto los verbos de percepción como toda la gama de características de los objetos significa tener sensaciones, por lo que, para él, el problema no consiste en averiguar cómo se modifican las reacciones espontáneas o aprendidas de los percipientes ante las características de los objetos cuando se alteran, dañan o lesionan los sentidos o el cerebro. Al metafísico le interesa otra cosa: dar cuenta de cómo los sentidos y el cerebro producen o dejan de producir "sensaciones" en los percipientes. Esto último nos revela, en primer lugar, que el término 'percepción'

evidentemente no tiene el mismo significado para el metafísico que para el científico y, en segundo lugar, que ese cambio de significado lleva al metafísico a investigar algo radicalmente distinto de lo que investiga el científico, a saber: la naturaleza de los procesos a través de los cuales los sentidos y el cerebro generan en los percipientes las supuestas representaciones mentales de los objetos físicos. Pero veamos rápidamente y en contraste lo que investiga el científico y lo que el metafísico quiere explicar con esos mismos datos. Voy a empezar con una breve semblanza de la versión científica del problema.

Al igual (supongo) que la mayoría de nosotros antes de ser contaminados por la teoría causal de la percepción, cualquier científico diría que los animales "perciben" cuando el científico en cuestión constata que, en condiciones normales, éstos son capaces de actuar y de responder adecuadamente a determinados estímulos que de antemano se sabe que los harán reaccionar de determinada manera. Esto hace que el término 'percepción' no sea un término simple, sino que normalmente exprese o describa la acción o reacción que los objetos, sonidos, sabores, olores y demás provocan en nosotros cuando se ven, se tocan, se oyen, se gustan o saborean, o se huelen u olfatean.

Partiendo de este modo de entender el término 'percepción', supongamos ahora que un científico tiene un gato que no sólo no reacciona ante ningún estímulo visual, sino que cada vez que intenta desplazarse de un lugar a otro, choca y se tropieza con cuanta cosa esté a su paso. Evidentemente, no se requiere ser científico para concluir que ese pobre gato tiene problemas visuales. Sin embargo, supongamos ahora que el científico descubre que su gato tiene una lesión cerebral y que después de una serie de estudios y comparaciones el científico descubre que cada vez que hay una lesión semejante a la que tiene su gato, los percipientes presentan problemas visuales semejantes a los que observa en su gato. La pregunta aquí es: ¿nos está descubriendo el científico algo acerca de la naturaleza de la visión al correlacionar lesiones cerebrales como las de su gato con problemas visuales? La respuesta es 'no', ya que el científico fue capaz de establecer la correlación entre lesiones cerebrales y problemas visuales sólo y porque el científico ya era capaz de distinguir cuándo un animal presentaba problemas visuales y cuando no, sin necesidad de explorar el cerebro de los percipientes. Lo que el científico sí descubrió fue que lesiones cerebrales como las de su gato normalmente van acompañadas con problemas conductuales catalogados como visuales. Pero éste descubrimiento lo pudo hacer gracias a que fue capaz de relacionar dos clases de hechos: los conductuales que le indicaron cuándo un percipiente tenía problemas visuales y los fisiológicos concernientes a la lesión cerebral. Los primeros son relevantes para la caracterización de la naturaleza de la visión, por lo que el descubrimiento de la lesión cerebral, aunque le proporcionó una información muy valiosa respecto a lo que podía estar asociado con una conducta visual errática, evidentemente no le reveló nada que él no supiera antes acerca de la naturaleza de la percepción visual - no sólo de su gato, sino de la percepción visual en general.

Esto es importante porque en este caso es la conducta del animal lo que nos permite decidir si éste es capaz o incapaz de ver, oír, etc. Empero, es obvio que para que el animal sea capaz de ejecutar dicha conducta requiere de mecanismos. Lo que el científico descubre al establecer correlaciones entre problemas conductuales y fisiológicos es que si los

mecanismos se alteran, dañan o extirpan, evidentemente el problema de los mecanismos se va a reflejar en la conducta del animal. En este caso, sin embargo, los mecanismos se relacionan con la conducta y no con las ideas o imágenes mentales que los metafísicos quieren suponer que dichos mecanismos generan.

El problema surge cuando el científico se deja enredar en el lenguaje del metafísico y cae en la fácil tentación de decir que cuando él descubre una lesión cerebral o cualquier otra clase de problema neurofisiológico relacionado con los denominados 'problemas visuales' lo que descubre son las áreas del cerebro dañadas que son responsables de producir la visión. En este caso es obvio que lo que se asume es que los elementos físicos asociados con los problemas visuales del animal son los responsables no sólo de ocasionar problemas conductuales visuales, sino de producir en los percipientes la visión de los objetos físicos. Pero esto, que a primera vista sólo parece un inocente juego de palabras no es, ya que, en este caso, el término 'visión' no refiere a una conducta, sino a la aprehensión mental de los objetos que tienen los percipientes. Esto lleva a suponer que la naturaleza no sólo de la visión, sino de la percepción en general debe quedar definida por los factores físicos que la teoría causal de la percepción asegura que intervienen, a saber, los objetos físicos, los órganos sensoriales y el cerebro.

Por otra parte, si a este malentendido le sumamos la idea de que existen dos clases de objetos, esto es, los físicos y los percibidos, y asumimos (como comúnmente lo hace el metafísico) que los objetos percibidos son causados por la acción que producen los objetos físicos: en nuestros sentidos y el cerebro, resulta entonces que parecería que para saber cuál es la naturaleza de la percepción de un animal no basta con observar sus acciones o reacciones ante los objetos cotidianos, sino que se requiere dar cuenta de toda la cadena causal que se supone que produce en ese animal "percepciones", las cuales, en lenguaje del metafísico quiere decir 'imágenes o representaciones internas de los objetos externos'. Pero veamos rápidamente como quiere dar cuenta el metafísico de esa cadena causal, concentrándonos una vez más en el caso de la visión.

El metafísico afirma que la visión debe ser entendida como el último eslabón de una cadena causal que se inicia en un complejo sistema de rayos u ondas luminosas, que son emitidos o reflejados por la superficie de los objetos y que "viajan" del objeto al ojo del percipiente. Cuando un haz de luz llega finalmente a la retina, se producen cambios químicos en los receptores, cambios que provocan que una serie de impulsos eléctricos viajen a lo largo de las fibras nerviosas hasta llegar hacia una de las dos áreas visuales del cerebro, que son las que reciben los impulsos nerviosos de la retina. Estos impulsos generan a su vez una intrincada actividad en esas áreas del cerebro así como en otras asociadas a ellas para, final y felizmente, provocar o causar en el percipiente la visión del objeto físico (Hirst, *Perception and the External World*, 1965, p.10).

Ahora bien, muchos de nosotros estamos más que familiarizados con este tipo de explicaciones y, desafortunadamente, por la manera como se usan en esta clase de discursos enunciados y términos que parecen provenir de la ciencia, la gente en general y, muy en particular los neurofisiólogos tienden a creer que esta cadena causal debe ser considerada

como un buen punto de partida para que efectivamente se logre ofrecer una explicación "científica" de cómo se producen en nosotros las percepciones de los objetos externos. Lo que no se considera es que esa cadena causal no sólo es un mito, sino que así como está elaborada termina enredando los temas de tal manera que se acaba por no saber ni lo que lo que los científicos afirman conocer ni cómo se pretende acceder empíricamente a una realidad distinta a la que se percibe. Veamos por qué digo esto.

Si queremos ser coherentes con lo implicado por la idea de una cadena causal como la mencionada, tendremos que admitir no sólo que existen dos mundos, el físico y el percibido, sino dos espacios: el exterior, donde se supone que habitan los objetos físicos, y el percibido, donde se supone que habitan los objetos que percibimos. Esto, de alguna manera nos obliga a suponer que el espacio percibido no es el externo, con lo cual nos vemos comprometidos con la idea de un espacio interno. Así, resulta que los objetos que percibimos tienen simultáneamente dos ubicaciones, a saber, la externa, donde se inicia la cadena causal y la interna, donde termina la cadena causal. Por supuesto que está la premisa de que el mundo que conocemos no es el físico sino el que percibimos, por lo que de alguna manera se tendrá que sostener que los rayos u ondas emitidos o reflejados por el objeto, así como el objeto mismo que emite o refleja la luz, aunque son las causas de nuestras percepciones no forman parte de nuestras percepciones. El resultado es que no nos es posible ver ni siquiera el ojo por medio del cual vemos. Este resultado nos lleva a concluir que, por definición, ningún enunciado expresado en el lenguaje físico es un enunciado empírico ya que, según lo establece esta misma cadena causal, lo que empíricamente percibimos no son los objetos físicos, sino el efecto que los objetos físicos (que incluye desde los rayos u ondas emitidos por los objetos hasta nuestros órganos de los sentidos y el cerebro) produce en el precipiente. Si esto es así, entonces tendremos que aceptar que todo lo físico que conforma la cadena causal - y que se supone explica los procesos que produce nuestras percepciones - no es corroborable empíricamente.

Pero los problemas e incongruencias no se detienen allí. El complicado proceso causal que se supone que tiene como función expresa dar cuenta de la manera como generan los sentidos y el cerebro nuestras percepciones del mundo externo a final de cuentas también se queda sin explicar: simplemente se asume que de algún modo misterioso el cerebro es capaz de traducir los cambios químicos que se producen en los receptores no sólo como impulsos eléctricos, sino como el conjunto de manchas coloreadas de determinada forma y tamaño, que se supone se ubican en nuestro espacio visual como las imágenes sensibles que funcionan para nosotros como las representaciones internas de los objetos externos.

Claro está que cuando se estudian los órganos de los sentidos y el cerebro con el instrumental científico pertinente (para, por ejemplo, medir su actividad), no es nunca posible detectar otra cosa que no sean cambios físico-químicos. No obstante, el metafísico normalmente asume que cuando los registros le indican que existen receptores, neuronas, vías cerebrales, etc., que se activan de cierta manera ante determinado estímulo, lo que se está indicando es qué receptores, neuronas, vías cerebrales, etc., están encargados de percibir ese estímulo. El metafísico, por supuesto, tiene que suponer que cuando se estudia

cómo se activan las neuronas, vías cerebrales, etc., ante cierta clase de estímulos (e.g. colores, formas, texturas, etc.), lo que se estudia es cómo perciben las neuronas o las vías cerebrales esos estímulos. Pero esto sólo puede ser visto así cuando *a priori* se asume que la percepción es algo que tiene lugar en el cerebro y que por 'percepción' debe entenderse algo que refiere a nuestras representaciones o imágenes internas causadas por la acción que producen los objetos externos en nuestros órganos de los sentidos y el cerebro.

El problema es que esta forma de entender el término 'percepción' no explica nada en la medida en que no está planteada como una hipótesis científica, sino sólo como una mera definición. Lo que esto quiere decir es, en pocas palabras, que la "explicación" causal acerca de longitudes de onda, receptores, neuronas, vías cerebrales, etc., no nos dice ni nos descubre nada acerca de la percepción. Lo que se hace es simplemente reflejar la manera como el metafísico quiere definir lo que debe entenderse por 'percepción' y todo el contenido "físico" del complicado proceso causal se arma *a posteriori* con base en su definición. La dificultad estriba en que, como ya vimos, por más que se investigue y se den más detalles acerca de cómo operan los órganos de los sentidos y el cerebro ante determinados estímulos, la definición que da origen a la cadena causal y la cadena causal misma no nos conduce a ningún lado.

No obstante y, a pesar de todo lo anterior, hay quienes defienden la validez de la explicación en términos de la cadena causal argumentando que es un hecho que, si rompemos cualquier eslabón de dicha cadena, nuestra percepción del mundo automáticamente se interrumpe. Lo que no se considera es que, aunque efectivamente poderos decir que nuestra percepción del mundo se altera o se cancela si los órganos de los sentidos o el cerebro se alteran o se dañan, la ciencia no tiene el poder y, de hecho, no debena de tener ningún interés en perder el tiempo tratando de "explicar" qué causa la percepción.

EMERGENCE: A TENSION BETWEEN DEFINITIONAL AND EMPIRICAL PROPERTIES

NYDIA LAIRA
Laboratorio de Neurocomputación
Centro de Instrumentos de la UNAM
Apartado Postal 70-186
México, 04510, D.F.
México

ABSTRACT

The aim of this paper is to elucidate two points. The first is that the concept of 'emergence' arise as an attempt to answer the following question: How is it possible for a whole to possess properties that are absent in its component parts? The second is that the idea that there are some qualities or properties that emerge from the complex game performed by small particles (defined by the physical sciences as devoid of secondary qualities) is the product of a conceptual confusion. This conceptual confusion arises when we take conceptual definitions, such as 'particles', to be the real constituents elements of the world, and with this conceptual platform pretend to explain the sources of those properties that we, human beings, inevitable perceive in our ordinary objects.

KEYWORDS: Conceptual definition, technical concepts, definitional properties, primary and secondary qualities, conceptual principles, ontological principles, science, metaphysics, materialism, vitalism, Emergentism.

INTRODUCTION

Conceptual definitions play an important role in the theoretical construction of any science because they allow us to build technical concepts. Technical concepts are definitions that fix in advance the limits of the meaningful use of a concept. Concepts are terms that usually describe objects or properties of objects (e.g., this is a table, this is red, this is alive, etc.). However, technical concepts, properly speaking, are not genuine concepts. They are what we might call 'formal concepts'. Formal concepts, in contrast with genuine concepts, do not describe objects or properties. They do not deal directly with objects, but with concepts, and they specify what has sense to predicate of concepts, not of objects. This characteristic allows technical concepts to incorporate in their conceptual definition the manner in which an object has to appear in propositions. Hence, objects within technical concepts are represented as mere variables that fall in a proposition under certain descriptions and no others.

Newton, for example, introduced in his *Principia* the technical concept of 'body' defining it as that which is extended, hard, impenetrable, movable, and endowed with its proper inertia. Through this definition Newton only pretended to give us a precise meaning of how exactly he was going to use the term 'physical body' in his mechanical theory. Given this definition, anything extended, hard, impenetrable, movable and endowed with its proper inertia could be treated as a mere variable that, within the frame of his theory, equally follows the same rules or laws of mechanics, whether it was an apple, a tree, a book, a table, an animal or a human being. Nevertheless, Newton did not pretend to say that what he was establishing to be the definitional properties of any physical body were the only properties we can use to describe bodies. Odor, color, taste, and many other qualities possessed by most objects we perceive daily were not included in his technical definition, not because he did not conceive them as real properties, but because they were unessential for his theoretical purposes. The color of an object, for example, was conceived by him as one of the many of the so called 'secondary qualities' because it was meaningless to consider the color of an object when the theory only pretended to offer those general rules and conditions which instruct us how to measure the movement or trajectory of what he defined to be a 'physical body'. He discards secondary qualities as well as others

because, in order to be able to announce those general rules that conform to the so called 'laws of mechanics', he did not need to take in consideration more properties in bodies than those that serve the purpose of his theory (A. Tomasini, *Los colores y su lenguaje*, p. 183-184).

Through this example I just want to say that technical concepts, even when they usually are derived from our daily concepts, do not pretend to include all the characteristics or properties we perceive or find in our daily world. Suffice, for the purpose of the mechanical theory, to endow bodies with those measurable properties that coincide with the general rules or conditions that might help us to define their motions, impacts, changes, etc. Color, odor, flavor and many other qualities are, strategically and by definition, taken away of the realm of the mechanical theory because those characteristics simply do not play any role in the mechanical explanation of bodily movements and changes. However, to exclude secondary qualities from the realm of a physical theory does not mean that those qualities are alien qualities in the real world. They are only purposeless for a mechanical account of movement and changes. And this is fixed in advance by the technical concepts that conform the theory.

COMPLEX AND SIMPLE BODIES

We, without doubt, perceive chairs, tables, pencils, human beings, apples, liquids, gases, etc. All these things are endowed with different qualities and accept different sorts of descriptions. Nevertheless, the analysis of the composition of our daily object in mechanical terms can be made without considering more qualities than primary ones. Lets us try to illustrate this through another example.

Newton notices that experience show us that most objects we perceive are susceptible to be divided or separated in lesser parts. This fact allows him to introduce the distinction between complex and simple bodies. Complex bodies were defined by him as bodies made of elementary physical particles, and physical particles were conceived to be like very small physical bodies. According with this distinction Newton sustains that

the extension, hardness, impenetrability, mobility, and inertia of the whole, result from the extension, hardness, impenetrability, mobility, and inertia of the parts; and hence we conclude the least particles of all bodies to be also all extended, and hard and impenetrable, and movable, and endowed with their proper inertia (Newton, *Principia*, Rule III).

Notice that physical particles as well as physical bodies are devoid, deliberately and by definition, of what for the purpose of the theory are to be considered as secondary qualities, and that the relation between wholes and parts is coherent with Newton's technical definition of 'body'. This coherence allows Newton, first, to treat wholes and parts as variables of the same class and, second, to build a theory with the power to unify in one and the same conceptual scheme the diversity of movements and changes we perceive in different objects according with its primary qualities only. However, 'particles' were introduced as conceptual entities (Newton calls them 'inferred' entities), and not as real entities. They are postulated in order to explain changes and movements in mechanical terms in what he previously defined to be a 'complex body' (including liquids or gases, which in our daily life do not fit with the conceptual definition of 'physical body').

Now, since Newton, physical sciences can adequately explain and describe the movements and changes of the object of the world solely in terms of primary qualities, and once we understand the physical theory in which the concept of 'complex body' and 'particle' apply, we also are in the position to understand the incredible power of physical sciences to manipulate, predict, or reproduce many changes and movements observed in the real world only in physical terms, that is, without considering secondary qualities in objects. Nevertheless, the explanatory power of physical sciences might produce a subtle conceptual change in our minds, and mistakenly take conceptual definitions, such as 'particles', to be the real constituents elements of the real world. When this happens we might confuse explanatory principles with ontological principles and truly believe that it is an incontrovertible fact that, for example, the whole realm of our perceived objects is

really made of physical particles devoid of secondary qualities. Before I go on, I think that some conceptual remarks are at stake here.

Ontological principles, in contrast with explanatory principles, are assumptions that can neither be proved nor disproved in any rigorous sense. These assumptions might have scientific grounds, but scientific results are taken to be facts of the real world, and not only coherent explanations of it. This means that, by a twist of our minds, we might be able to assume ontological compromises with conceptual entities, and take for granted that, for example, physical particles, devoid among other things of secondary qualities, have to be understood as the real component entities of the world we live in. Philosophically speaking, this assumption conforms one of the most important conceptual supports of those metaphysical doctrines known as 'materialism'.

MATERIALISM

In modern times most forms of materialism have a strong link with physical sciences. However, materialism is not a scientific theory, but a metaphysical doctrine. As a doctrine, materialism does not restrict itself to define and delineate the general conditions of how things might move, link or change according with the technical definitions and laws that the theory provides. It wants to account for those causes or agents that make things move, link or change the way they do in our daily life. This means that, in contrast with the modest procedure that conform a scientific theory, metaphysical doctrines want to go behind the scene stipulated by a physical science, and tell us not only what the world is really made of, but also how things become to be what they are.

As a metaphysical doctrine, materialism sustains that everything is made of matter. Matter is believed to be made of elementary material particles, and these material particles are conceived to be, not just explanatory principles, but the primitive substance of the real world. According with most materialism doctrines, all the variety of things that we are able to perceive in our daily world is a material thing. And every material thing is made up of different arrangements of elementary particles. Elementary particles are conceived to be units of matter that have various fundamental forms, such as protons, mesons, electrons, positrons, neutrons and many more. These particles can be endowed with mechanical, electrical and chemical properties that allow them to move, link and interact with each other in space and time according with certain laws (e.g., physical, chemical, electrical). It is also held that the arrangements and patterns of behavior of particles conform the objects we perceive daily. Hence, particles endowed with those primary qualities defined by scientists; and their arrangements and behavior in space and time form a complete ontology of the real world, and they are all that seems to be required in order to explain the source and causes of the whole realm of all the macroscopic objects we perceive daily.

Now, the link between materialism with physical science is more than obvious. However, an important conceptual change is at stake here. Materialism sustains that macroscopic objects have to be understood as the effect of the internal arrangements of particles. This entails not only the ontological assumption that the internal arrangements of particles have to be understood as the causal source of all macroscopic objects, but also the epistemic assumption that it is possible to reconstruct the whole world of macroscopic objects through the knowledge of its internal structure. Both assumptions conform the basis of what has been call the 'mechanistic' paradigm.

MECHANISTIC PARADIGM

Methodologically speaking, the mechanistic paradigm entails its own way to study complex systems and bodies. It proposes that in order to understand the behavior of a system as a whole it is necessary: a) to dismantle the system in its component parts, b) to analyze the behavior of each isolated part, and, c) to reassemble the system trying to understand how the behavior of each element is affected when combined with other elements.

cómo se activan las neuronas, vías cerebrales, etc., ante cierta clase de estímulos (e.g., colores, formas, texturas, etc.), lo que se estudia es cómo perciben las neuronas o las vías cerebrales esos estímulos. Pero esto sólo puede ser visto así cuando *a priori* se asume que la percepción es algo que tiene lugar en el cerebro y que por 'percepción' debe entenderse algo que refiere a nuestras representaciones o imágenes internas causadas por la acción que producen los objetos externos en nuestros órganos de los sentidos y el cerebro.

El problema es que esta forma de entender el término 'percepción' no explica nada en la medida en que no está planteada como una hipótesis científica, sino sólo como una mera definición. Lo que esto quiere decir es, en pocas palabras, que la "explicación" causal acerca de longitudes de onda, receptores, neuronas, vías cerebrales, etc., no nos dice ni nos descubre nada acerca de la percepción. Lo que se hace es simplemente reflejar la manera como el metafísico quiere definir lo que debe entenderse por 'percepción' y todo el contenido "físico" del complicado proceso causal se arma *a posteriori* con base en su definición. La dificultad estriba en que, como ya vimos, por más que se investigue y se den más detalles acerca de cómo operan los órganos de los sentidos y el cerebro ante determinados estímulos, la definición que da origen a la cadena causal y la cadena causal misma no nos conduce a ningún lado.

No obstante y, a pesar de todo lo anterior, hay quienes defienden la validez de la explicación en términos de la cadena causal argumentando que es un hecho que, si rompemos cualquier eslabón de dicha cadena, nuestra percepción del mundo automáticamente se interrumpe. Lo que no se considera es que, aunque efectivamente poderos decir que nuestra percepción del mundo se altera o se cancela si los órganos de los sentidos o el cerebro se alteran o se dañan, la ciencia no tiene el poder y, de hecho, no debena de tener ningún interés en perder el tiempo tratando de "explicar" qué causa la percepción.

emergentism is, without doubt, a form of materialism in the sense that it accepts that everything is made of matter, and also that all material things are made of elementary particles. According with other forms of materialism, it also accepts that nothing happens without some motion of elementary particles. However, emergentism has two peculiarities that distinguish it from other materialism doctrines. The first is the fact that, in contrast with other forms of materialism, emergentism has not antecedents in the history of philosophy. We can trace the first steps of emergentism only in the middle of the nineteenth century with John Stuart Mill, Alexander Bain, and G.H. Lewes. The second, closely related with the first, but indubitably much more important, is that the source of emergentism is strongly linked with the somehow predictable inconsistencies that can be found in the exaggerated metaphysical explanatory power that some metaphysical theories wanted to ascribe to the mechanistic paradigm.

It is important to bear in mind that metaphysical doctrines, even when they try to support their philosophical positions on scientific grounds, their assertions are not restricted to observational scientific reports. Metaphysical doctrines usually become cosmological constructions that want to give an account of the universe as a whole. And the best that can be done when a doctrine wants to extend scientific data and its (restricted, by definition, to specific realms of events) to the whole universe, is to give plausible arguments on the grounds of adequacy. When a metaphysical doctrine is puzzled by inconsistencies, the doctrine is forced to revise its arguments and principles in a fundamental way. In that case science is useless and only the internal coherence of the doctrine can save it. Let us try to illustrate what is at stake with some historical remarks.

Biologists and materialists found that an ontology based solely in the existence of material particles, devoid of any other things, of secondary qualities, life, consciousness, purposiveness, etc., were more than inconsistent when they have to explain the presence of characteristics and properties absent, by definition, in material particles, but ostensibly present in many of our daily objects. The inconsistencies of some of the basic assumptions of materialism were clearly detected when the characteristics of living systems (e.g., growth, digestion, procreation, as well as goal directed and the so called 'intelligent' behavior) wanted to be explained by the kind and arrangement of material parts and the general laws of physics that hold for all things that conform all the objects in the world (A. Beckermann, *Reductive and Nonreductive Physicalism*, 1992). This attempt is labeled by biologists under the name 'reductionistic' or 'mechanistic' program. However, and despite the undeniable and overwhelming results of the reductionistic program to interpret biological phenomena only in physical terms, biologists traditionally have found a strong discomfort to the idea that life is nothing more than a complex interplay of a large number of particles. The reiterated resurgence of vitalism doctrines can be seen as an expression of the distress felt by biologists when living systems become an aggregate of lifeless physico-chemical elements. However, most forms of vitalism only rebel about the ontological principles of materialism. They usually do not put in question the analytical coherence of the mechanistic paradigm, neither the fact that this method is the logical result of a complex interplay between a theory and a body of evidence that depends on the coherence of those technical definitions among other characteristics, deliberately exclude the so called 'vital' phenomena from its realm of

traditionally those metaphysical doctrines that are puzzled not only with life, but with the so called 'mind-body problem', as well as with secondary qualities, want to solve conceptual confusions as if they were the result of empirical lack of data. During the XIX century, for example, vitalism try to solve the "mystery" of the so called 'vital' phenomena, postulating the existence of an extra element: a certain nonmaterial element, a substance-like *entelechy* or *elan vital*. This extra element mistakenly was taken to be a specific hypothesis that, on the grounds of "scientific" research, biologists could detect. In time, however, this "extra substance" simply was declared to be a "false hypothesis" because nobody was able to find more instances in living bodies than those equally possessed by nonliving systems. Without noticing that something that was defined to be nonmaterial could not be accepted or refuted in empirical grounds, biologists keep on trying to solve the conceptual distress caused by the idea that living systems have to be distinguished in something "extra" to distinguish them from other bodies. But neither reductionism nor vitalism was able to solve the "mystery" of the presence of life and other properties and qualities observed in some bodies

without finding strong inconsistencies in their "reconstruction" not only of life but of our most familiar objects guided by the mechanistic paradigm.

EMERGENTISM

I mention some of the quarrels that took place between materialism and vitalism for two reasons: the first is to illustrate that conceptual misunderstandings cannot be solved by empirical research. The second is to explain that right in the middle of the biological battle between reductionism and vitalists' positions, emergentism finds its way in the history of materialism. Although, emergentism cannot be considered as a common form of materialism.

Emergentism wants to adjust some of the complaints stressed by vitalists concerning the impossibility to understand living systems characteristics through the knowledge of its parts, but preserving, at the same time, the ontological assumption that the whole universe is made of physical particles endowed only with primary qualities. To reconcile both ideas, emergentism has to put in jeopardy the epistemic assumption of the mechanistic paradigm that asserts that it is possible to obtain a complete knowledge of wholes through the knowledge of its component parts. This epistemic assumption is closely linked with the old cosmological idea that "no causes can give rise to products different from those of the same class."

Now, science and metaphysics have to different versions with respect of the relation that has to be established between wholes and parts. Science infers the qualities, regularities, movements, and structures of parts according with the qualities, regularities, movements and structures of wholes. This means that, even when scientists in principle do not contradict the causal principle, they do not use it. They use, instead, what we might call the 'analogy principle' that ensure them that parts obey the same rules that we empirically find that wholes obey in the frame of a theory. There are no mysteries in science because scientists (when they do not get involve in metaphysical speculations) do not have to explain what causes what.

If we recover Newton's definition of 'particle', clearly he derived from the definitional properties of wholes the definitional properties of parts in order to deal with parts and wholes as conforming the same class of variables that obey the very same rules of movement. Parts are inferred from wholes, but parts do not play a causal role in the theory. Nevertheless, when we want to interpret, as materialism does, that parts are the causal source of wholes, and keep the principle that asserts that there cannot be more in the effect than there is in the cause, some conceptual inconsistencies arise when we want to derive the properties we perceive in wholes through the properties that, by definition, are absent in their component parts.

Newton was able to exclude many observed qualities from his technical definition stipulating that his theory was not the study of matter, but the study of its movements. Through this stipulation he clearly separate metaphysical considerations about the properties of matter from science, and simply proceed to define wholes and parts of bodies according with his aims. Secondary qualities were not mysterious qualities that science has to explain as effects of their parts. They were only qualities that did not play any role in a theory of movement. However, metaphysicians want to derive the properties of wholes assuming the existence of parts that only posses those properties that had been chosen to explain their movements. Obviously, in the metaphysical relation between wholes and parts the presence of secondary qualities and the like, appear to be a mystery to be explained. The aim of the concept of 'emergence' is to explain that "mystery". And an example given by C.D. Broad might help us to illustrate the point. I quote Broad:

Take an ordinary statement, such as we find in chemistry books; e.g., "Nitrogen and Hydrogen combine when an electric discharge is passed through a mixture of the two. The resulting compound contains three atoms of Hydrogen to one of Nitrogen; it is a gas readily soluble in water, and possessed of a pungent and characteristic smell." If the mechanistic theory be true - and an archangel with unlimited mathematical competence could be able to know all microstructures that can be realized - the archangel could deduce from his knowledge of the microscopic structure of atoms all these facts but the last. He would know exactly what the microscopic structure of ammonia must be; but he would be totally unable to predict that a substance with this structure must

smell as ammonia does when it gets into the human nose. The utmost that he could predict on this subject would be that certain changes would take place in the mucous membrane, the olfactory nerves and so on. But he could not possibly know that these changes would be accompanied by the appearance of a smell in general or the peculiar smell of ammonia in particular, unless someone told him so or he had smelled it for himself. If the existence of the so-called 'secondary qualities', or the fact of their appearance, depends on the microscopic movements and arrangements of material particles which do not have these qualities themselves, then the laws of this dependence are certainly of the emergent type. (Broad, *The mind and its Place in Nature*, p.71-72).

Now, the so called 'emergent laws' represent not only the abandonment of the causal principle, but an appeal to inexplicable factors. Emergentism justifies the intervention of inexplicable emergent factors introducing the biological notion of 'evolution'. The peculiar interpretation of this notion allows emergentism to assert that, as well as higher organisms can emerge from lower ones, there must be a hierarchy of levels of organizational complexity of material particles that includes, in ascending order, the strictly physical, the chemical, the biological, and the psychological level. From here the name 'Emergent Evolutionism'.

If emergent evolutionism is or is not the "solution" of some of the metaphysical controversies around the presence of those qualities and properties that we perceive daily without any distress, is a matter to try to understand what kind of explanation is that which have to introduce the work of mysterious emergent factors in order to save the inconsistencies of materialism. As I point out before, only when metaphysicians want to derive the properties of wholes assuming that parts are their source (containing only the characteristics that scientists stipulate in their technical definitions), our daily perception of colors, taste, smells, etc., become qualities which sources have to be explained.

Emergentism wants to say that water, for example, has the emergent characteristic of translucency, or ammonia the emergent characteristic of a pungent smell on the grounds that these characteristics cannot be deduce from the complete knowledge of the microstructure of those atoms that conform water or ammonia. However, it is obvious that neither translucency or the pungent smell is derived from any set of statements about their chemical compound. But it is unnecessary to introduce the work of mysterious emergent factors to explain them. Suffice to understand that the terms 'translucency' and 'pungent smell' simply are vacuous when what we want is to define the chemical compound of water or ammonia. Furthermore, 'translucency' and 'pungent smell' are terms that we use to describe some characteristics of water or ammonia. But, oxygen, hydrogen, nitrogen, and the like, do not work as descriptions of some characteristics of water or ammonia. They are definitions, and their role in our language is to fix the limits, in advance, of how water or ammonia has to be conceived in order to define them in chemical terms.

BIBLIOGRAPHY

- BECKERMANN, FLOHR and KIM, *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*, Walter de Gruyter, Berlin, N.Y., 1992.
 BROAD, C.D., *The Mind and its Place in Nature*, Routledge and Kegan Paul, London, 1925.
 NEWTON, I., *Principia*, Vol. II, Motte's Translatio, Revised by Cajori, University of California Press, USA, 1996.
 TOMASINI, A., *Los Colores y su Lenguaje*. Anthology, Percepción: Colores, Instituto de Investigaciones Filosóficas, México, 1993.

PROBLEMAS CONCEPTUALES Y METODOLÓGICOS EN TORNO A LA NOCIÓN DE 'SISTEMA'

Nydia Lara Zavala
Laboratorio de Neurocomputación
Centro de Instrumentos, U.N.A.M.
Apartado Postal 70-188 Coyacán, 04510
México, D. F.

RESUMEN

Actualmente gran parte del trabajo desarrollado en la interfase que connota el estudio de los seres vivos con el de modelos computacionales se hace recurriendo a una metodología que podríamos quizá denominar de "sistemas". Esta metodología sin duda permite formularse y estudiar en entornos múltiples, tareas con comportamiento (externo e interno) de los seres vivos. Sin embargo, pocos han reparado en los problemas de interpretación que contiene esta metodología. Lo que queremos explorar en este trabajo es cómo, a partir de los postulados que encierra la noción de "sistema", se logra fusionar el estudio de los seres vivos con la construcción de máquinas artificiales. Parte de nuestro objetivo es mostrar que la noción de "sistema" no surge, como se tiende a pensar, de una teoría científica, sino de una metafísica de corte "organicista", la cual, como trataremos de explicar a continuación, al combinar el estudio de los seres vivos con el formalismo que ofrece la lógica-matemática desarrollado de la tecnología computacional, logra conformar el punto de vista de acuerdo con el cual se sostiene que cualquier sistema puede ser estudiado a través de otro sistema. Ésta señala las mismas propiedades "funcionales" o dinámicas internas. Lo que nos interesa a nosotros revisar en esta ocasión es que el enfoque organicista a pesar de que indudablemente se ha podido aplicar exitosamente a múltiples disciplinas, tiene la desventaja de que acarrea consigo la idea de que la función o tarea que realiza un sistema la determina la misma dinámica interna del sistema. Por esto, como lo iremos viendo poco a poco, es un error que acarrea consigo problemas que son muchos los que se refieren a la filosofía de la mente que a los que le compete a la ciencia explorar.

BIOLOGÍA TEÓRICA, ORGANICISMO Y TEORÍA GENERAL DE SISTEMAS

El organicismo, como lo mencionamos anteriormente, debe ser interpretado como una doctrina metafísica y, como tal, cabe señalar que sus propuestas no son susceptibles de ser rechazadas por la vía científica, lo que nos permite afirmar que, como que propugna esta doctrina, si bien no directamente los resultados del trabajo de investigación, si la interpretación que se hace de esos resultados, por lo que vale la pena iniciar el análisis de la metodología que utilizan actualmente todas aquellas disciplinas que trabajan con el enfoque de "sistemas" reduciendo brevemente las ideas más importantes que conforman al organicismo.

El término 'organicismo' fue introducido técnicamente en la literatura por L. von Bertalanffy, a quien se le considera el padre más de la teoría general de sistemas, sino del movimiento que se conoce con el nombre de "biología teórica". El organicismo, tal como se concibe en la actualidad, surge como una propuesta que podríamos catalogar como híbrida en la medida en que se estructura con el fin de reconciliar dos visiones que en su momento se consideraron como mutuamente excluyentes en torno a las características esenciales que conforman a los seres vivos. Nos referimos, en concreto, a lo que los biólogos teóricos denominaron "vitalismo", por un lado, y a lo que denominaron "materialismo" o "reduccionismo", por el otro. Digamos unas cuantas palabras sobre lo que según ellos sostiene cada una de estas posturas para que se entienda lo que se supone aporta el organicismo a la concepción del ser vivo.

El "vitalismo" es el nombre que se le da a una vieja doctrina filosófica que tiene como su gran exponente a Aristóteles. Sin embargo, no es la doctrina aristotélica la que discute el biólogo teórico, sino la versión que se introduce como franca oposición a los incipientes esfuerzos de la biología molecular por reducir el estudio de los seres vivos al estudio de sus componentes físico-químicos. En oposición al "reduccionismo" de la biología molecular, el vitalista sostiene que un ser vivo no se puede estudiar de la misma forma que se estudia una compleja pieza de maquinaria. Desde su perspectiva es simplemente imposible entender el funcionamiento de un ser vivo analizando el papel que juega cada una de sus partes en la maquinaria. La prueba de ello es que al probar de que el estudio de los componentes físico-químicos que conforman un ser vivo requiere de un instrumental científico muy preciso para generar sus resultados, las interpretaciones que se generan

de este enfoque rara vez pueden ser integrados en un todo coherente. De hecho, las aplicaciones del ser vivo a nivel molecular son constantemente rechazadas de piezas nuevas de un rompecabezas que siempre es incompleto. Esto tiene como efecto el que el biólogo molecular tienda a pensar que lo que le hace falta es la biología son nuevos experimentos, para generar nuevos datos. Pero, para el vitalista, el fracaso del enfoque molecular para explicar la vida, su génesis, su desarrollo y su comportamiento o actividad, radica en el hecho de que desamblar al ser vivo para entender su funcionamiento equivale a eliminar exactamente aquello que distingue al ser vivo del resto de la materia.

Desde la perspectiva del biólogo vitalista, aunque es cierto que el ser vivo está compuesto de materia inorgánica, la actividad que lo distingue del resto de la materia tiene que responder al hecho de que en lo vivo existe otra clase de sustancia, la cual no puede ser concebida como física ya que es esta sustancia la que le imparte a lo físico poderes que no existen si no se trata de un ser vivo. Sin embargo, también existe una diferencia esencial entre la materia viva y la inerte, la cual se expresa en la capacidad que tiene "lo vivo" por regular o dirigir la actividad de sus componentes físicos hacia un fin determinado. El aspecto finalista del ser vivo no puede ser ni causada ni explicada por medio de las leyes, conceptos, métodos que ofrecen las ciencias físicas, puesto que son las causas no físicas las que en última instancia determinan la peculiar actividad que manifiestan los cuerpos físicos cuando están vivos.

Ahora bien, aunque el biólogo teórico se nutre de la crítica que lleva a cabo el vitalismo del enfoque molecular, la tesis organista enfáticamente rechaza la idea de que se necesite postular la existencia de una entidad no física para dar cuenta de las características peculiares que distinguen a los seres vivos de los seres vivos. En esta perspectiva, podemos decir que el organista acepta la tesis que el biólogo molecular, que el ser vivo está compuesto en su totalidad de elementos físico-químicos. Sin embargo, el organista, al igual que el vitalista, sostiene que el enfoque molecular, por sí mismo, no sirve para comprender el ser vivo. La razón que ofrece el organista es, empero, muy distinta a la propuesta sustancialista que ofrecen los vitalistas, ya que, para los primeros, el fracaso del enfoque molecular no responde, como sostienen los segundos, a que "lo vivo" contenga una sustancia extra distinta a la materia física, sino a un problema que es que el enfoque molecular, que el biólogo teórico normalmente empleado por el biólogo molecular. El problema consistía en que el biólogo molecular piensa que el conocimiento del funcionamiento de cada parte que compone a los seres vivos, tarde o temprano darán como fruto el conocimiento del funcionamiento del todo orgánico. Pero así proceder, si bien es cierto que nos da mucha información sobre los elementos que contiene un organismo, irremediablemente el biólogo molecular en un mar de datos y resultados que acaban por no decirnos nada acerca del funcionamiento del organismo. El error del biólogo molecular, para el organista, consiste en ignorar que el ser vivo funciona como una totalidad, que cada una de sus partes refleja la relación que ésta tienen con las otras partes que conforman al todo. De hecho, la tesis que sustenta al organicismo como tal consiste en afirmar que el ser vivo tiene que ser comprendido como una unidad de organización compleja formado por distintos niveles de organización que se sobrepone en forma de manera jerárquica. Cada nivel de organización tiene sus propias características y sus propias leyes, por lo que cada nivel puede ser analizado por sí mismo, siempre y cuando se comprenda que algunas características del funcionamiento de algún nivel de organización está determinado por el funcionamiento de otro nivel de organización. El organista acepta la idea de que la única manera de dar cuenta de los aspectos de regulación y dirección que caracteriza la actividad de los seres vivos es mediante un punto de vista teleológico, aspecto que podía ser incluido en la visión "organicista" de los seres vivos. Esto lo que quiere decir en pocas palabras, es que para el organista el ser vivo no puede ser entendido como un todo, lo que determina el funcionamiento de sus partes. Tesis que de factor permite el método *zoom-up* seguido por el biólogo molecular y lo reemplaza por una aproximación *top-down*, donde primero se determina la función del sistema y luego se analiza cómo lo hace el sistema. Este enfoque o forma de tratar el ser vivo continúa como una vieja doctrina que, a pesar de que arranque del método que caracteriza el enfoque de sistemas que abre el camino que uno los intereses de la biología teórica con el estudio de sistemas artificiales. Veamos, pues, brevemente cómo es que el teórico de sistemas logra establecer la conexión entre el estudio de los seres vivos y el diseño y construcción de sistemas artificiales.

El teórico de sistemas trata de reunir dentro del enfoque *top-down* dos perspectivas diferentes relacionadas con un mismo objeto de estudio, *viz*, el comportamiento de los seres vivos. Las dos clases de estudio tienen que ver con lo que en la jerga ingenieril se denominan respectivamente "descripciones externas" y "descripciones internas" de un

sistema. La primera aproximación ofrece una descripción funcional del comportamiento del ser vivo, lo cual implica que para dar cuenta de lo que hace el ser vivo, previamente tiene que haber una interpretación semántica de lo que significa la acción del ser vivo. La segunda clase de aproximación, en cambio, consiste en el examen de la acción que ejecuta el ser vivo con miras no a detectar y analizar el ensamblado anatómico y fisiológico de aquellos mecanismos que suponemos que participan de manera esencial en el desarrollo de esa acción. Naturalmente se trata de dos clases de estudios totalmente diferentes y, como lo mencionamos, dan origen a dos tipos de descripciones que son radicalmente pero completo diferentes. Pero sobre este punto hablaremos un poco más adelante; por lo pronto lo que nos interesa explorar es cómo es que el teórico de sistemas logra reconciliar estas dos clases de descripciones que se puede sostener, sin temor a equivocarnos, que son lógicas y fácticamente independientes entre sí.

Empecamos por aclarar que el teórico de sistemas reconoce que hay una diferencia entre las descripciones externas y las internas, pero sostiene que dicha diferencia no es significativa si consideramos dos clases: La primera, que se refiere a descripciones que describen diferentes clases de actividades de manera simultánea; la segunda tiene que ver con la idea de que dentro del organismo existen diferentes niveles de organización jerárquica. Esto último implica que se requieren diferentes descripciones para dar cuenta de las diferentes actividades de un sistema, de manera que se entiendan en términos de partes (o cosas) relacionadas entre sí para realizar una función definida.

Ahora bien, recordamos que detrás de la noción de "sistema" está la organización jerárquica, tesis que sustenta la idea de que el comportamiento del organismo depende de la interacción y del funcionamiento de las partes que lo conforman. El estudio del comportamiento del ser vivo normalmente requiere de la previa observación de ciertas patrones de regularidad. El reconocimiento de que la acción tiene una función no lleva de manera natural a encontrar un significado a la acción o a dicha función. La descripción de la actividad del organismo un inevitable aspecto metodológico. Es esto último lo que le permite al teórico de sistemas conectar las descripciones externas con las internas a través de una aproximación top-down, la cual parte de las descripciones externas que determinan qué hace fundamentalmente el sistema, para luego proceder a buscar las descripciones internas necesarias como para que nos permita entender la manera como internamente se lleva a cabo esa función en los distintos niveles jerárquicos de su organización. Dado que la acción de un nivel de descripción interna está subordinado a la función que se define en el nivel de descripción superior, cada descripción siempre llevan una carga semántica que les da a las descripciones internas significado conductual. Por eso, las descripciones internas en muchos casos dan la apariencia de que explican el proceso interno que lleva a cabo el sistema para poder hacer lo que hace. De aquí la idea de que lo que hace internamente el sistema puede ser genéricamente entendido como un procesamiento de información. Es esta noción la que permite ligar el estudio de los seres vivos con las máquinas (computadoras) procesadoras de información y es, sin duda, lo que reúne al teórico de sistemas con la cibernética y posteriormente con el conjunto de teorías científicas actualmente en boga que se refieren con el nombre de "ciencias cognitivas". La clave para conjugar los intereses del biólogo, con los de los matemáticos, los ingenieros, los psicólogos, etc., es, por supuesto el concepto de "información", por lo que vale la pena detenernos un poco en él, ya que, como veremos, su definición acarrea una ambigüedad que establece una "información" a pesar de que matemáticamente está perfectamente armado.

INFORMACIÓN Y PROCESAMIENTO DE INFORMACIÓN

Posiblemente le debemos a N. Wiener la idea de que los problemas relacionados con la comunicación y el control tanto en seres vivos como en máquinas, podían ser estudiados con total independencia del contexto de donde surgían. El control del azúcar en la sangre, los movimientos que realiza un animal al capturar su presa, al igual que el control de la temperatura de un calentador o de la presión de un reactor, podían ser abordados con las herramientas formales que ofrecían la teoría de la comunicación y la ingeniería de control. Estas dos áreas conjuntamente eran vistas como para dar cuenta de aquellos problemas que se relacionan con la comunicación y el control, con el fin de involucrar las nociones de "regulabilidad" y "direccionalidad" que tanto inquietaba a los vicietistas. Después de todo, los problemas relacionados con la regulabilidad y la direccionalidad son problemas estrechamente ligados con la comunicación y control de mensajes, y todos los procesos de control, es decir, cuando se trata de los procesos de flujo de información que lleva las "instrucciones" que determinan el comportamiento global del sistema en sus distintos niveles jerárquicos. Estas "instrucciones" se conciben como mensajes y la cantidad de información que transmiten estos mensajes (señales), está por supuesto

relacionada con la cantidad de información que puede ser transmitida a través de lo que se denomina un "canal de información". Shannon y Weaver se ocuparon, como todo el mundo sabe, del problema de transmisión de mensajes, pero reconocieron que, aunque su teoría logró generar un modelo de los sistemas de comunicación muy versátil en sus aplicaciones, por la manera como está concebido no contempla el valor semántico de la información. La medida del contenido informacional que contiene un mensaje, no tiene nada que ver con lo que el mensaje significa. No obstante, Wiener utilizó el término "cantidad de información" como sinónimo de "cantidad de significado" que se refiere a la información que Shannon claramente establece entre tres aspectos distintos relacionados con el término "información". El primero, según Shannon, tiene que ver con la solución del problema técnico del que se ocupa la teoría de la información, el cual está relacionado sólo con la capacidad de un sistema para transmitir información. El segundo, refiere a la capacidad que el receptor, lo cual, evidentemente, no es susceptible de formalización.

Ahora bien, si recordamos que el planteamiento de los modelos organizistas se sustentan en la idea de que es posible conectar las descripciones que se refieren a las intenciones a través de la intencional carga semántica que les confiere a las descripciones internas significado conductual, es evidente que se pueda pensar que estos modelos guardan alguna relación con el nivel tres que reconoce Shannon. Pero esto, como lo tratáramos de ejemplificar a continuación, sólo puede hacerse cuando se conforma una red de transmisión de información. En la medida en que se utiliza una aproximación top-down asumiendo como válido el enfoque de los teóricos de sistemas.

Un ejemplo clásico de cómo opera el método top-down lo podemos encontrar en la propuesta de la teoría de esquemas de M. Arbib. Este autor sostiene que se puede describir una conducta específica, como puede ser la de reconocimiento de patrones (e.g., discriminación entre presa-predador). Esta conducta Arbib la considera como una "caja negra" que puede ser decompuesta en otras cajas más pequeñas; captura de información, fijación, evitación, cambios de atención, intersección dar cuenta de la función de la caja más grande (i.e., el reconocimiento de patrones). La misma estrategia se continúa completando cada una de las cajas pequeñas en cajas aún más pequeñas. Así, por ejemplo, si tomamos la caja de fijación de la presa, la podemos decomponer en orientación, aproximación, fijación binocular, lengüetazo, enguilmiento y limpieza. Esta decomposición, según lo determina la teoría de esquemas, se puede continuar hasta que el proceso de información definido por las cajas más pequeñas pueda ser explicado por una red neuronal" (ver Fig. 1). Viendo la figura, uno pensaría que salvó la red neuronal, todas las descripciones internas de los organismos son ambiguas, sin embargo esto es muy ambiguo en la teoría de esquemas, ya que Arbib postula que las cajas representan procesos de información en jerarquías más altas. Esto parece justificar la idea de Arbib de que para que la acción de un sistema con respecto a una conducta específica, los organismos requieren tanto de "esquemas perceptuales" como "motores", elementos que, según él, dan cuenta de la habilidad del sistema para realizar esta diversidad de tareas. Así, tenemos que la conducta de captura de presa se interpreta como una caja que tiene que ver con reconocimiento de personas, reacciones de escape, etc. que está relacionado con intenciones que se refieren a una conducta específica que analizamos la descripción conductual que se hace de esta tarea. Esta ambigüedad no es un problema particular de la teoría de esquemas de Arbib, se trata más bien de un problema intrínseco al enfoque de sistemas, ya que según el teórico de sistemas se requieren diferentes descripciones para dar cuenta de las diferentes actividades de un sistema, de manera que cada descripción interna debe tener su contraparte en una externa.

Lo que es importante tomar en cuenta para evitar el tipo de ambigüedad al que se enfrentó con la teoría de esquemas es que las descripciones externas, como correctamente lo reconoce el teórico de sistemas, son funcionales, lo cual quiere decir que estas son sólo interpretaciones adecuadas para dar cuenta de lo que hace un sistema. Esto, en pocas palabras, significa que las descripciones externas nunca pueden dar cuenta de las intenciones internas de un sistema, ya que se refiere a la interpretación que nosotros los hablantes le adjudicamos a esa acción. Así, cuando decimos, por ejemplo, que un tigre caza porque tiene hambre, por incomprensión de cómo opera nuestro propio lenguaje tendemos a pensar que cuando utilizamos el término "hambre" describimos un estado del tigre, cuando en realidad "hambre" es el nombre que utilizamos para darle significado a la acción del tigre. Esto nos lleva a la siguiente consideración: no hay propiamente hablando una cosa tal como la fisiología de la conducta o la descripción interna de la conducta, simple y sencillamente porque no hay una cosa tal como la

fisiología del significado ni una descripción interna que de cuenta del significado.

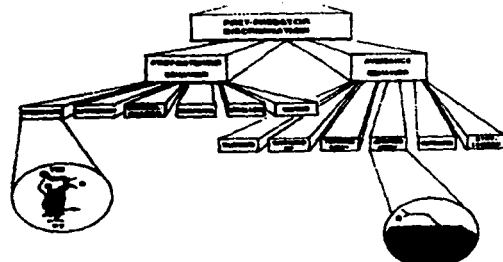


Figura 1
Coordinación visuo-motora en anfibios según la teoría de Esquemas.

Ahora bien, aunque de hecho se nos puede establecer correlaciones entre las descripciones externas e internas, lo que nosotros queremos señalar es que estas últimas sólo nos pueden brindar información en tanto y en cuanto la actividad interna de un sistema bajo estudio cuando lo exponemos a determinado estímulo y exploramos sus patrones de activación, digamos, con las técnicas propias de las neurociencias. Lo que podemos determinar en este caso es sólo qué y cómo qué dinámica se activan los distintos elementos del SNC al animal. Sin embargo, la dinámica que se observa y que puede servir como guía para generar tanto modelos numéricos susceptibles de ser simulados por una computadora o por un modelo electrónico, son significativas, esto es, si bien hay un correlato de afuera hacia dentro, no se puede establecer si el mismo correlato se mantiene de dentro hacia afuera.

ON ANALYSIS OF CORRESPONDENCES

L.A. MENDOZA

DIVISION DE ESTUDIOS DE POSGRADO
FACULTAD DE INGENIERIA

RESUMEN Este es el primero de dos artículos en el que se introduce un nuevo enfoque de convergencia de sucesiones generalizadas de subconjuntos y de similitud de correspondencias (funciones multivaluadas). Se comparó este nuevo enfoque con definiciones y resultados publicados en la literatura, y se muestra que este nuevo enfoque es una generalización de estos últimos. Se introduce el concepto nuevo de conjunto de acumulación de sucesiones generalizadas de subconjuntos. En el segundo artículo se presentan algunas aplicaciones e sistemas de control y a optimización.

1. INTRODUCCION

La formalización matemática de correspondencias (funciones multivaluadas) comenzó cuando desde principios de siglo, si no es que desde finales del siglo pasado (Poincaré, Kuratowski [4]). Las aplicaciones de la teoría son muchas más recientes, el mayor lo son las aplicaciones en control de sistemas determinísticos (antico espacio de estado).

Los resultados que se presentan en estos artículos tratan de subsanar el hecho de que se han estado utilizando resultados teóricos sobre correspondencias en control de sistemas determinísticos, sin que se tenga todavía una teoría sólida de correspondencias (en el sentido de completa y general), y esto al analizar lo conocido, mientras las definiciones sean aplicables sólo a casos particulares está presente la posibilidad de obtener contradicciones en la teoría sin advertir la operación de éstas.

Este es el primero de dos artículos acompañantes, y en éste se establecen resultados teóricos, mientras que en el segundo se habla de aplicaciones a control de sistemas determinísticos. En estas líneas se hace un análisis formal de la noción de convergencia de una sucesión de subconjuntos y de la definición de consistencia de una correspondencia, incluyendo el caso de un espacio topológico (en el contexto más general posible, que es el de espacios topológicos). Se comparó con algunas definiciones y resultados publicados en la literatura, se explica porqué estos últimos constituyen casos particulares, y se dan ejemplos que hacen patente la necesidad de generalizar

dichas definiciones y resultados particulares (aplicaciones a mencionarse en el artículo acompañante).

La mayor parte de este introducción se concentra en una discusión conceptual. En la siguiente sección se dan algunos resultados técnicos, pero no se incluyen las demostraciones matemáticas (para esos datos puede consultarse el informe de investigación [5]).

Prólogo. Dada una sucesión de subconjuntos de R^n (o de cualquier espacio topológico) se tienen las siguientes dos alternativas:

- (i) sucesión de subconjuntos como una sucesión de puntos en el espacio potencia de R^n .
- (ii) considerarlos tal cual está: como una sucesión de subconjuntos de R^n .

Con respecto a la primera alternativa, se debe topologizar el conjunto potencia de R^n para poder hablar de convergencia de una sucesión (esto es, dotar al espacio potencia de una norma, métrica, o una topología en el caso general).

Cuando los elementos de la sucesión son subconjuntos compactos, se puede utilizar la métrica de R^n para definir una métrica en la colección X de todos los subconjuntos compactos de R^n (colección que es un subconjunto propio del espacio potencia de R^n). Desde entonces, se aplica la métrica de Hausdorff (ver [1]), y hay de modo sencillo una noción de convergencia de sucesiones de subconjuntos compactos inducida por la métrica de Hausdorff.

Ahora bien, utilizando la noción de convergencia usual de sucesiones de puntos de R^n , se puede definir una noción de convergencia en la colección de todos los subconjuntos cerrados de R^n (que tienen un subconjunto propio del espacio potencia de R^n). Los trabajos de Poincaré y Kuratowski [4] dan el siguiente estado de la teoría. Dicha definición de convergencia generaliza la mencionada ya, a través de la métrica de Hausdorff.

En caso alguno se utiliza la segunda alternativa para definir convergencia de una sucesión de subconjuntos $\{A_n / n \in N\}$, sea $A_n \subset R^n$ conjuntos arbitrarios (i.e. no necesariamente cerrados) en R^n . En dicha definición en el contexto más general de sucesiones generalizadas de subconjuntos en espacios topológicos. Esto último nos permite ver la colección de imágenes de un conjunto cerrado de R^n . Desde entonces, se aplica la métrica de Hausdorff (ver [1]), y hay de modo sencillo una noción de convergencia de sucesiones de subconjuntos generalizada (F, [1, 1]; R^n). Ejemplos de correspondencias objeto de estudio de teoría de control de sistemas determinísticos se mencionan en el artículo acompañante.

Cabe enfatizar que la segunda alternativa que se escogió puede arrojar más luz sobre lo que sucede en las aplicaciones que la primera alternativa, por ejemplo, si se habla del conjunto de estados alcanzables de un sistema de dichos subconjuntos es no vacía, relación que no es obvia si se está trabajando en el espacio potencia, en este último espacio simplemente se sentirá que dichos conjuntos de R^n son dos puntos x y y en el espacio potencia de R^n tales que la distancia entre ellos es estrictamente mayor que cero (y se que $x \neq y$).

Nesta época es de nuestro conocimiento, no hay resultados publicados en la literatura para el caso de sucesiones de subconjuntos en el sentido general mencionado en el párrafo anterior.

Otro punto sobre el que cabe hacer observaciones en esta introducción es el de conjunto de acumulación. Sobre este concepto no hay mención alguna en la literatura publicada, en el presente momento, supiéramos que se observa la evolución en el tiempo de un fenómeno que requiere para su modelado en cada instante de varios parámetros y supiéramos que éstos dependían del tiempo. Supiéramos que el fenómeno que se está estudiando es un subconjunto de números reales que incluyen los parámetros mencionados y mediciones de las variables de estado, se tendrían construidas de modo natural, las subsecciones $\{A_n(n) / n \in N\}$ para cada instante. Cada una de ellas es un conjunto de números reales, se puede hablar de su convergencia, (su conjunto límite debe ser dicho conjunto constante).

Esta subsección sería convergente, en que por ello la sucesión de la que fue extraída tenga que ser convergente. Ahora bien, dicha subsección fue extraída de la sucesión original a través de un criterio bien definido, que no es satisfecho por todos los elementos de la sucesión original.

Cabe incluir más observaciones que revisan el ejemplo anterior y la segunda alternativa (ii) mencionada en la página 2. El límite inferior de la colección de subconjuntos (definición de Poincaré-Kuratowski [4]) utiliza los puntos de acumulación de sucesiones de puntos, cada punto pertenece a alguno de los subconjuntos que conforman la sucesión de subconjuntos, sin necesidad de hacer la elección de Poincaré-Kuratowski [4] en el siguiente, desde que tales subsecciones de puntos no se contemplan si los conjuntos que incluyen los subsecciones de puntos son los mismos, para que correspondan a los mismos subconjuntos, de modo que no hay posibilidad alguna de distinguir entre diferentes subsecciones de subconjuntos y los diversos conjuntos de acumulación asociados a éstos.

Se hace notar también lo siguiente: dada una sucesión constante de subconjuntos, $A_n = A$, $n \in N$, con A un subconjunto propio abierto de un espacio métrico X (esto es, existe $x \in X$ tal que $dist(x, A) > 0$ y tal que $x \in A$), cualquier definición de convergencia que se proponga debe de ser que x no

SYSTEMS SCIENCE AND THE STUDY OF COMPLEX BEHAVIOR

NYDIA LARA-ZAVALA

**Laboratorio de Neurocomputación
Centro de Instrumentos
Universidad Autónoma de México
Apartado Postal 70-186
México, 04510, D.F.**

KEYWORDS: Systems science, cause, purposive behavior, teleological and mechanical descriptions, tasks and movements.

ABSTRACT

The aim of this work is to show that System Science has a good method to understand the so called 'purposive behavior', nevertheless, people have some conceptual misunderstandings of their own method, must of all because they are introducing conceptual problems inside their research.

In this work I am going to explain the basic characteristics of the method employed to understand the complexity that distinguish the analysis of purposive behavior, and afterwards I am going to explain how conceptual problems are introduced in the description of the mechanisms of the system in order to learn to avoid them.

INTRODUCTION

Under the name of 'Systems Science' it is possible to find the participation of a constellation of disciplines that come from a variety of fields that belong to Engineering, Physics, Mathematics, Biology, and Humanities. All these disciplines have in common the interest to understand the internal mechanisms that conform the behavior of the so called 'complex' or 'organized' systems. Characteristically features of those systems are:

- a) to be formed of multiple elements
- b) those elements interact with each other
- c) the interactions between elements are, in general, non linear
- d) the activity of one element is affected and affects the activity of the other elements
- e) the activity of each element appears to contribute to fulfill the performance of the global behavior of the whole system, and
- f) the joint behavior of all the elements of the system seems to fulfill certain functions, goals, or tasks.

It is the aim of this work to analyze why these two last characteristics are introducing conceptual problems to properly understand the method that has been used to explain the so called 'purposive behavior' of complex systems.

Let us begin remembering that the sorts of mechanisms that explain the behavior of a complex system begin to be seriously examined at the middle of this century, in particular by those biologists and engineers that were puzzled about how to model purposive behavior. Two brand-new theoretical fields raised from this effort. The first was the so called 'theoretical biology' started by L. von Bertalanffy. The second was started by N. Wiener, and conform what we might call 'theoretical control engineering', or, more precisely, the 'abstract theory of controlled artifacts'¹. The work of Bertalanffy culminated in what he calls 'General System Theory'; the work of Wiener in what he calls 'Cybernetics'. The link of both seminal works conforms today the field of interest of Systems Science. Nevertheless, and despite its name, the main concern of Systems Science is not to conform a new discipline, but rather to solve some of the theoretical, conceptual, and methodological puzzles that arise when people want to explain purposive or teleological behavior in mechanical terms. Let us try to understand what is behind this idea.

TELEOLOGICAL AND MECHANISTICAL BEHAVIOR

One of the most original contribution of Systems Science is the method that help us to study the mechanisms that underlie the so called 'purposive behavior'. Unfortunately, the authors as well as users of this method do not realize that in order to understand how the internal mechanisms of the system under study participate in the fulfillment of a specific function, goal or task, it was necessary to define, at least, two different and logically independent items: the first is the task to be performed by the system, lets say, to reach a glass of water; the second is the description of the movements, lets say of the arm, that the system performs to fulfill that task (e.g., ballistic movement, finger adjustment, hand rotation (Arbib, 1989)).

This division entails two great advantages: one is that if we accept that there are no logical connection between the task and the different movements that the system performs to fulfill it, in principle, it is acceptable to assert that any system (natural or artificial) apt to perform the same or similar movements in the right sequence, could be trained, instructed or programmed to execute the succession of movements required to fulfill that task. The other is that the different movements that the system perform can be associated either, with the fulfillment of the task, or with the search of the mechanisms that the system has to have in order to be able to perform the required movements (see fig. 1).

Thus, the very same movements that a system performs to accomplish a task, can be seen as teleological, if we relate the whole sequence of the movement of the arm with the fulfillment of that task, or as mechanical, if what we want is to understand how the different muscles,

¹ Engineering is the technological discipline that designs and builds artifacts. These are artificial systems that fulfills a given function to solve a certain problem.

tendons, nerves, or brain activities, participate and interact one with the others when the arm performs those movements.

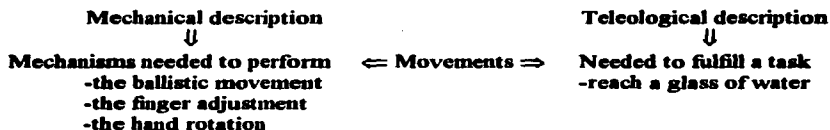


Fig. 1. It is possible to explain the very same movements in teleological terms if what we want is to understand what is achieved with them, or in mechanical terms if what we want is to understand what mechanisms are needed to perform them, independent of the task.

EXTERNAL AND INTERNAL BEHAVIOR

It is important to bear in mind that no movement, by itself, has a purpose. This is so because, in the first place, the purpose is defined by the task, not by the movements the system perform to reach it, and, in the second place, because the purpose is never an extra movement or a cause for other movement (even when we usually employ that term suggesting it). Purposes, well understood, are only our way to describe what is possible to achieve when some movements are performed in the right context, and nothing else is implied when we talk about purposes. This semantic reflection is important because now we are in the position to separate the purpose of the task from the movements the system has to perform. Doing this, it is possible to proceed to study those movements in terms of the design of those internal mechanisms that permit the system to perform them (either in context or out of context). At this point, what really cares is to find out the proper way to include or understand the elements and their interactions that the system needs to perform the movements it requires in the right sequence.

Obviously, this last item demands from us to consider at least two different sequences of behavior produced by the same system at the same time. The different behavioral sequences are: the external, that the system (the arm in this case) performs as a whole, and the internal (the muscles contractions, the role of the tendons, the nerve transmissions of signals to the brain, tendons and muscles, the brain activity, etc.), that has to be performed by those elements that conform the mechanisms that the system has to have in order to carry out the expected behavior (in this case, the ballistic movement, the finger adjustment and the hand rotation).

Both behavioral sequences (the external and the internal) are closely related with the fulfillment of the task. Nevertheless, the aim of the research is not the task that the system has to perform, but the understanding of the design of the internal mechanisms that might allow a system to perform the external sequences of movements we need it to perform. At

this stage we can leave aside the purpose of the behavior to concentrate our attention in how the system is or has to be built in order for it to be able to execute those external movements mechanically, that is, without a purpose. This means that Systems scientists are not interested in the study of specific systems performing specific tasks, but rather in knowing the internal mechanisms that the system needs in order to perform the right external movements in the right sequence.

Note that the conceptual distinction between the purpose of the task and the study or design of the mechanisms that the system needs to perform the movements required to fulfill it, is conceivable because, as we remark before, the purpose is defined by the task, not by the movements the system performs. This last issue is important for two reasons: the first is that the concern of the research is now concentrated in the understanding of the structure and dynamics of the mechanisms, and this research have to do no more considerations than to find out what mechanisms are needed to obtain the right movements in the right sequence. The second is that, at the systemic structural level, it is not important to take into account if the system is or is going to be made of neurons, chips, or whatever material we can think of. What matters at this level is the possibility to obtain the structure and dynamics of those mechanisms that will allow the system to perform all the movements we expect to be performed mechanically. This obviously entails that the analysis and synthesis of the structure and dynamic of the mechanisms is not related with the nature or characteristics of the system and its component parts, but only with the possibility or impossibility to obtain the most complete knowledge of the mechanisms that a system requires to be able to perform the right movements in the right sequence. Nothing more is important at this stage.

These feature has many conveniences. One of them is that engineers, for example, can use, as inspiration, the knowledge of the mechanisms possessed by those living systems that already are capable to perform the movements they want to be performed by their machines. And in the same fashion, engineers can help biologists to discriminate, in the complex architecture that usually characterize living systems, which elements and interactions are relevant and which are not in the performance of certain movements. In time, the richness of this interaction has proved that one and the same structure of a general mechanisms can help both, engineers and biologists, in the sense that engineers can use that general structure to design and build their machines, whereas the biologist can use the same design as an aid to identify the mechanisms possessed by their systems. One historical example of this approach was the so called 'negative feedback' conceived by Wiener. It was used to build a thermostatic control with the ability to turn a unit off or on whenever certain temperature level were exceeded, and the very same design was exported to biology for the understanding of one of the most recurrent mechanisms employed by living systems to control not only its temperature, level of sugar, etc., but many other state variables.

Another advantage is that, the prospect to distinguish between the task and the movements performed by the system, has given the opportunity to link the study of the design of mechanisms with other research fields. Sociologists, economists and politicians, for example, had found in this procedure two interesting notions to understand their systems. One of

them is the richness that is behind the idea that systems of different nature can be design to produce exactly the same sequence of movements or activities. The other one is the idea that, if there is no logical connections between the definition of a task and the movements that the system performs mechanically, it is possible to know what task we want a system to be able to perform without knowing which movements the system has to perform to reach the expected function, goal or task. Both ideas together has given the opportunity to all those scientists that usually have a clear idea about what task they want their systems to fulfill, to use computer simulation to explore which possible movements can be performed by their systems and, through computer simulations obtain the knowledge of the final results that their systems could achieve when they perform certain movements and avoid others.

METHODOLOGICAL CONSIDERATIONS

So far, we have been able to appreciate how a series of conceptual distinctions (*e.g.*, task and movements, purposes, internal mechanisms, etc.) allow Systems Sciences to offer a method not only capable to obtain clear empirical results, but also capable to favor the interdisciplinary work between different disciplines. The improvement of this method consists in the fact that the concern of the research is centered simultaneously in two different items: one is the prediction, reproduction or simulation of those external movements that might allow a system to perform a task, the other is the effort to find out the design of the mechanisms that might allow a system to perform those external movements.

The examination of the movements or activities that a system performs to fulfill a task allow us to obtain two different comprehension of the very same movement or sequence of movements: one teleological, the other mechanical. The first describes what the system obtains when it performs those movements. The second describes the elements and interactions that are needed to perform those movements.

Now, usually when we design the structure and dynamic of the mechanisms that a system needs to perform certain movements, we have to keep in mind not only which movements we want the system to perform, but which task we expect to be performed through those movements. This seems to tolerate a combination of teleological and mechanical descriptions. Nevertheless, if we forget that mechanical descriptions not only are different but independent of teleological descriptions, and subtly mix the description of the activity of the mechanisms with the purpose of the action, shortcuts might begin to replace the clarity of the method and the task can be understood as forming part of the mechanisms.

This is what happens with those people that, for example, following the method step by step, find the correct design of the mechanisms that allow their systems to obtain the sequence of movements that a system requires to fulfill certain task, and afterwards want to know how pure physical elements and interactions are able to determine internally what to do to fulfill that task. The main difficulty of this conceptual shortcut is that it seems to be enough to produce a disturbing conceptual inaccuracy. This inaccuracy consists in thinking that if a

mechanism is needed to perform a movement, and that movement is required to fulfill certain task, then, the elements and interactions that conform the mechanism not only has to be understood as the causal source of the movement, but as the causal source of the purpose that those movement fulfill. Note that through this conceptual twist, mechanisms now appear not only suitable to produce a specific movement, but suitable to produce it with specific purposes. Purposes, then, begin to form part of the mechanism. Idea which is not only totally absurd but logically contradictory if we remember that, by definition, mechanisms are only another level of description of one and the same movements, and movements by themselves have no purposes.

Nevertheless, when people stop thinking in the meaning of their terms and are willing to ascribe purposes to mechanisms, the empirical research of those mechanisms begin to be disturbed by metaphysical considerations. One of them is that the analysis of the internal activity of the mechanisms, induce them to believe that the coordinated activity they observe in each one of the elements that conform the system, is an activity that is being guided by a mysterious force. The main difficulty of this idea is that when people explore the elements that conform the mechanisms in order to find out the source of that mysterious force, they find nothing that allow them to explain where this mysterious force comes from. At this point science fiction tends to replace the lack of data obtained by the empirical research and, in addition, alien entities such as emergent properties, souls, minds, *elan vitals*, *entelechies*, etc., begin to form part of the explanation concerning the source of that force that nobody is able to find. Nevertheless, and as we try to explain in this work, that mysterious force finds its place in history only when people conceptually mix their teleological with their mechanical descriptions of those external movements they observe a system is capable to perform, curiously, sometimes with and sometimes without a purpose.

BIBLIOGRAPHY

- Arbib, M., *Visuomotor Coordination: Neural Models and Perceptual Robotics*, *Visuomotor Coordination: Amphibians, Comparisons, Model, and Robots*, Edited by Jörg-Peter Ewert & Michael A. Arbib, Plenum Press, N. Y. & London, 1989.
- System Behaviour*, Edited by John Beshon and Geoff Peters, The Open University Press by Harper & Row, Publishers, Great Britain, 1972.
- Klir, G. J., *Facets of System Science*, Vol. 7, Plenum Press, N. Y., 1991.
- Kamps, G., *Self-Modifying Systems in Biology and Cognitive Science*, Pergamos Press, Great Britain, 1991.
- Rossen, R., *Anticipatory Systems, Philosophical, Mathematical & Methodological foundations*, Vol. 1, Pergamon Press, Great Britain, 1985.

CONCLUSION GENERAL

No cabe duda de que durante los siglos XVI y XVII, la crítica abierta, la revisión conceptual y el análisis de la coherencia, o incoherencia, de un discurso con pretensiones de ser catalogado como científico, ayudó enormemente a la conformación de lo que hoy conocemos como la física newtoniana. Cuando Newton inició el arduo camino que lo conduciría a separar a su física de la especulación metafísica, la Royal Society y sus *Philosophical Transactions*, para poner sólo un claro ejemplo, jugaron un papel importante en la consolidación del proceso que sin duda culminó en uno de los más espectaculares logros del intelecto humano. Filósofos, matemáticos y científicos intervenían cotidianamente y al unísono en la revisión conceptual, formal y experimental de los aciertos y/o desaciertos contenidos en el discurso científico. No obstante, y aún en vida de Newton, el sano esfuerzo que consiste en separar, hasta donde sea posible, el discurso científico de la especulación metafísica quedó, desde entonces, en el olvido.

Este olvido sin duda ha propiciado el resurgimiento de la metafísica en la ciencia y con ello la búsqueda de causas metafísicas que sólo distraen la labor científica al tratar de dar respuestas a pseudo-preguntas que no esclarecen nada, pero sí complican la comprensión de lo que se debe buscar. Causas metafísicas, por ejemplo, empañan la relación entre el estudio de los seres vivos y la síntesis de robots. Prueba de ellos es la idea de que parte del quehacer de un investigador consiste en ubicar cómo y dónde se llevan a cabo procesos mentales tanto en los seres vivos como en las máquinas denominadas 'inteligentes'. Cerebros y máquinas con capacidades sensoriales se postulan para "explicar" la relación entre movimientos físicos y la realización de una función, meta, o tarea concretas, pero dichas explicaciones, para no decir pseudo-explicaciones, no sólo no elucidan nada, sino que se vuelven tierra fértil de interminables discusiones sin sentido.

Ahora bien, el objetivo que motivó la elaboración de cada uno de los doce trabajos que se recopilaron en esta tesis, fue el de poner en claro el hecho de que, cuando menos en este momento de su historia, la biología, junto con la gama de disciplinas que actualmente analizan y modelan el comportamiento propositivo, están más enfocadas a la especulación metafísica que en el trabajo serio y esclarecedor que caracteriza a la investigación científica. Ejemplos de la inmensa gama de discusiones que la metafísica genera alrededor del estudio de este particular comportamiento son los trabajos *Methodological considerations in Cognitive Science*, *Methodological considerations in System and Cognitive Sciences*, *Conductismo y psicología cognitiva*, *Mente, cuerpo y ciencias cognitivas*, *Emergence: a tension between definitional and empirical properties* y *Problemas conceptuales y metodológicos en torno a la noción de 'sistema'*. A través de estos trabajos espero haber dejado en claro que son varias las razones que han propiciado los enredos teóricos, conceptuales y metodológicos. Dentro de ellos podemos destacar los siguientes:

- a) El determinismo de Laplace, erróneamente atribuido a Newton, junto con el mecanicismo que de él se deriva, ha contribuido de manera certera a generar la dañina confusión en torno a la concepción que la física newtoniana ofrece de lo

que va a ser definido como materia y la concepción metafísica de la materia. Esta confusión es la que da pie a que se piense que la descripción completa de la materia que conforma a un sistema y el conocimiento de sus interacciones son condiciones necesarias y suficientes para explicar y predecir el comportamiento del sistema.

- b) Esta idea, junto con la tendencia del metafísico de explicar las causas que dan origen al universo, a la vida, a la percepción, al pensamiento, al lenguaje, o a la inteligencia, ha generado una lucha encarnizada entre materialistas (mecanicistas o reduccionistas) y vitalistas (donde se incluyen el vitalismo ingenuo, el organicismo y el emergentismo) que evidentemente no conduce a ningún lado.
- c) Las confusiones metafísicas que se desprenden de (a) y (b) están llevando a los científicos a tratar de encontrar mecanismos de procesos inexistentes (e.g., los procesos vitales o mentales que se postulan como el origen de comportamientos propositivos).
- d) La imposibilidad de encontrar los mecanismos responsables de los denominados 'procesos' vitales o mentales, también está llevando a los investigadores a suponer que aún les falta información sobre su sistema, o a adjudicarles a componentes o subsistemas específicos (e.g., moléculas, neuronas, cerebros, computadoras, etc.) capacidades vitales o mentales (e.g., moléculas con capacidad de leer, neuronas con capacidad de percibir, cerebros con capacidad de pensar, imaginar, o sentir, máquinas inteligentes, etc.).

El esclarecimiento teórico, conceptual y metodológico desarrollado a lo largo de esta tesis, nos permite concluir que, cuando menos algunos de los problemas que están distrayendo a los científicos con consideraciones metafísicas se pueden evitar reconociendo que:

- A) El estudio del comportamiento propositivo, como se explica en los trabajos titulados *Forms of explanation* y en *System science and the study of complex behavior*, requiere de más de una forma explicativa para lograr entenderlo y modelarlo. Sin embargo, hay que evitar tratar de reducir un nivel de explicación a otro. Con esto queremos decir que, aunque es necesario considerar la finalidad o el resultado al que lleva, o puede llevar, el comportamiento de determinado sistema o clases de sistemas, ni se debe buscar, ni se puede encontrar dicha finalidad estudiando los elementos e interacciones que componen al sistema. La razón es muy simple: la finalidad del comportamiento no es un atributo o característica que provenga del sistema, sino que dicha finalidad refiere a lo que nosotros conceptualmente calificamos como logro o fracaso del comportamiento que exhibe el sistema.
- B) Las únicas formas explicativas que deben de evitarse en la ciencia son las explicaciones teleológicas y las mecánicas (entendiendo por teleológicas aquellas explicaciones que intentan poner el fin del comportamiento como su causa y por mecánicas a aquellas que intentan explicar las causas teleológicas a través de los elementos y mecanismos que componen a un sistema). La razón para evitarlas es obvia: al científico no le compete dar cuenta del origen causal de nada. Su labor

debe limitarse a entender los patrones de regularidad, o irregularidad, del comportamiento que exhiben los sistemas que estudia. Dicha comprensión puede darse a muchos niveles de análisis (e.g., social, conductual, orgánico, molecular, etc.), pero ningún nivel de análisis le va a revelar nunca lo que causa u origina el comportamiento que estudia. No obstante, y como se muestra en casi todos los trabajos presentados es esta tesis, son éstas formas explicativas las que normalmente emplean los investigadores precisamente interesados en entender y modelar el comportamiento propositivo.

Basta considerar los puntos (A) y (B) al momento de elaborar los reportes de los trabajos científicos, para poder eliminar:

- 1) la intromisión de procesos metafísicos para explicar el comportamiento propositivo, ya que los procesos vitales o mentales sólo aparecen cuando el investigador intenta dar cuenta de lo que causa o genera la finalidad del comportamiento del sistema;
- 2) la infructuosa búsqueda de propiedades, cualidades o mecanismos responsables de generar el comportamiento propositivo, ya que esta búsqueda es sólo una de las consecuencias que se derivan del punto (1);
- 3) la sensación de que aún falta explicar algo más para acabar de entender cómo funcionan los sistemas que realizan una función, meta o tarea concretas, a pesar de que, en ocasiones, el investigador ya cuenta con: a) el conocimiento del comportamiento del sistema, b) la idea clara de los resultados de ese comportamiento, c) la descripción detallada de los elementos que conforman al sistema y d) el registro exacto de los mecanismos que emplea el sistema para ejecutar los movimientos que requiere su comportamiento.

Evidentemente, cada uno de los puntos aquí mencionados tiene consecuencias que afectan seriamente muchas de las interpretaciones de los datos y resultados que nos están ofreciendo diversos investigadores. Algunas de esas interpretaciones no sólo están plagadas de metafísica, sino que arrastran inconsistencias e incoherencias que es necesario no sólo reconocer y evitar, sino erradicarlas definitivamente del discurso científico.

En esta tesis, por ejemplo, se trabajó la relación entre la teoría causal de la percepción y las neurociencias. En el trabajo titulado *Metafísica, procesos metales y observaciones empíricas*, creo que logramos dejar claro que investigadores de reconocido prestigio están utilizando la teoría causal de la percepción para interpretar los datos y resultados de sus experimentos. Tres de mis trabajos tratan de explicar esta teoría con el objeto de resaltar sólo algunos de los gravísimos problemas, inconsistencias e incoherencias que ella conlleva. Aunque las repeticiones de algunos pasajes fueron inevitables, creo que en todos ellos se puede apreciar, en primer lugar, que lo que le da cuerpo a esa teoría es la idea de que la percepción debe ser entendida como un efecto de la realidad externa y, en segundo lugar, que la aceptación de esa idea no sólo confunde las descripciones físicas con las metafísicas, sino que inevitablemente nos conduce a plantear enigmas sin solución (e.g., el problema mente-cuerpo, la existencia de otras mentes, el problema relacionado con la

posibilidad de acceder al conocimiento del mundo externo, el problema de la significación del lenguaje, el problema del solipsismo, etc.).

Ahora bien, es claro que, cualquier investigador que arme sus experimentos siguiendo los lineamientos estipulados *a priori* por la teoría causal de la percepción, por un lado, va a tratar de interpretar sus datos y resultados experimentales según lo estipula la teoría y, en segundo lugar, el investigador se va a ver en la necesidad de involucrar todos y cada uno de los problemas, inconsistencias e incoherencias que la teoría contiene. Esto evidentemente sistemáticamente va a impedir no sólo la posibilidad de que los datos y resultados obtenidos por los neurofisiólogos puedan quedar armados en un todo coherente, sino que también va a impedir que a través de esos datos y resultados podamos obtener una idea clara en torno a las verdaderas funciones del cerebro y el papel que juegan sus mecanismos en lo que refiere a la ejecución normal o anormal de la conducta del animal.

Con base en nuestro trabajo, podemos concluir que la teoría causal de la percepción ha logrado que muchos neurofisiólogos carezcan de una visión acertada, y fuera de discusión, respecto a la función (o funciones) que debemos adjudicarle al cerebro. Algunos de ellos se sienten satisfechos otorgándole *a priori* funciones cognitivas; pero otros, quizá más sensatos, les cuesta trabajo aceptar que un pedazo de materia, por bien organizada que esté, sea capaz de llevar a cabo funciones que evidentemente no son físicas. Empero, y hasta donde yo he podido rastrear la bibliografía, rara vez el segundo grupo cuestiona los fundamentos que llevan al investigador a otorgarle a ese particular órgano funciones que directa o indirectamente tienen que ver con los denominados 'procesos mentales'. Como consecuencia de este descuido tenemos que ningún dato o resultado extraído de las interpretaciones científicas sirven para confirmar o rechazar la idea de que es necesario postular la existencia de un cerebro o una mente pensantes. Esto se debe, en parte, al hecho de que la gente no duda de que uno de los objetivos de la ciencia es explicar las causas que generan los procesos cognitivos y, en parte, porque detrás de esta búsqueda está la idea de que todo lo que hace o es capaz de hacer el sistema, en principio es explicable en términos de algún proceso que acontece al interior del sistema. Si a esto le agregamos la distinción que establece la teoría causal de la percepción entre realidad externa y realidad percibida, obtenemos que el esquema que utilizan los neurofisiólogos para estudiar la actividad del cerebro, prejuzga que lo que debe de hacer el cerebro es, primero, captar, procesar, entender, etc., la realidad externa, para después definir el, o los movimientos, que internamente necesita llevar a cabo el sistema para obtener los resultados deseados por el cerebro del sistema. Es importante notar que bajo esta interpretación, aunque se supone que el cerebro lleva a cabo dos clases distintas de procesamientos: los mentales y los físicos, en todo caso la causa de los movimientos físicos que observamos, tiene que quedar adjudicada a alguna clase de proceso mental. Esto es así porque se supone que los movimientos físicos que ejecuta el sistema tienen inmersa la función, meta o tarea que determina el cerebro o la mente del mismo sistema.

Para finalizar, quiero enfatizar que con los trabajos presentados en esta tesis quizá no podamos aún desprender a la biología de toda su carga metafísica, no obstante espero haber contribuido, aunque sea un poco, a esclarecer que la metafísica inmersa en las neurociencias

no sólo está generando confusiones respecto a las funciones del cerebro en relación con las capacidades motrices de los sistemas que lo poseen, sino que dichas confusiones impiden la posibilidad de entender qué tanto se ha logrado avanzar en la comprensión del funcionamiento de ese órgano. Dado este primer paso, ahora cuando menos estamos en la posición de afirmar que una permanente revisión teórica, conceptual y metodológica, cuando menos en algunas de las áreas de las neurociencias, más que un tema interesante, es una labor que debe considerarse no sólo necesaria sino urgente.