

EPISTEMOLOGÍA E HISTORIA DE LA CIENCIA

SELECCIÓN DE TRABAJOS DE LAS XV JORNADAS

VOLUMEN 11 (2005)

TOMO II

Horacio Faas

Aarón Saal

Marisa Velasco

Editores



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



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



La simulación computacional de los procesos mentales ¿revela la realidad?

Andrés Martín*

Introducción

La mayor de las preguntas que giran en torno de la inteligencia artificial puede escribirse como ¿cuánto de realidad hay en los modelos computacionales?, pregunta que se volvería más urgente en caso de alcanzar alguna simulación exitosa de la mente y sus procesos psicológicos.

No es casual el hecho de que la mayoría de los ataques hacia esta disciplina apunten a la posibilidad "lógica" o "en principio", de conseguir dicha simulación.

Nosotros no vamos a considerar esto último, sino que nos quedaremos con la primera pregunta: ¿cuánto de realidad hay en los modelos computacionales?

Resulta evidente el movimiento del pensamiento que fundamenta la importancia de la pregunta: si podemos programar una máquina de tal suerte que simule *adecuadamente* un fenómeno cualquiera, entonces debemos conocer alguna *verdad* (de la misma naturaleza que las *verdades* cotidianas) acerca de lo que estamos simulando.

Tal vez sea nuestra *actitud ontológica natural* (como dice Fine) la que nos conduzca a esta conclusión. Pero también podemos brindar una *respuesta realista* o una *antirrealista* a la pregunta. En el primer caso, diríamos que existe alguna suerte de *correspondencia* entre los conocimientos involucrados en el modelo computacional y los sucesos reales del acontecimiento. En el segundo, diríamos que la simulación se adecua al mundo, pero que este hecho nada nos dice acerca de cómo es el mundo real.

Me parece interesante abordar la cuestión de la simulación computacional de los procesos mentales dentro del marco de la polémica entre realistas y antirrealistas porque estimo que puede proporcionarnos valiosos aportes para el análisis de los límites y alcances de las investigaciones desarrolladas a través de estas metodologías.

Valga la aclaración de que en este trabajo no voy a intentar resolver las cuestiones que enfrentan a los defensores del realismo con sus antagonistas antirrealistas. Sólo pretendo utilizar sus posiciones para analizar algunas cuestiones relacionadas con una forma de desarrollar conocimientos. Embarcado en este objetivo, seguiré principalmente dos líneas argumentativas, la elaborada por Bas van Fraassen en *La Imagen Científica* y la desarrollada por Nancy Cartwright en *How the Laws of Physics Lie*.

Una buena simulación ¿y algo más?

Comencemos considerando que tenemos una simulación ya funcionando de los procesos mentales.

Preguntemos en primer lugar cómo hacemos para determinar el éxito o fracaso de dicha simulación.

* Universidad Nacional de Tucumán.

Creo que en este punto no caben mejores criterios que aquel establecido, hace tiempo ya, por Turing, es decir su famoso test: el simulacro y el suceso real deben ser indistinguibles. Que, para ir acercándonos al terreno de la discusión, puede ser traducido de manera precisa en términos *vanfraasceanos* como "el simulacro debe adecuarse empíricamente a lo que es el caso". Un ejemplo imaginario clásico sería una máquina con la que pudiéramos mantener una conversación sin darnos cuenta de que efectivamente nuestro interlocutor es una máquina.

Spongamos que efectivamente se satisface este criterio de adecuación ¿podemos concluir entonces que analizando el modo como hicimos la simulación obtendremos una lista de "todo lo que hay en nuestra mente"? O, dicho de otro modo, ¿debemos creer en la verdad de las teorías asumidas en nuestra programación del simulacro gracias a que éste es adecuado empíricamente?

Desde el antirrealismo postulado por van Fraassen evidentemente la respuesta es: ¡No! van Fraassen sostiene en la enunciación de su *empirismo constructivo*: "La aceptación de una teoría involucra como creencia solamente que ella es empíricamente adecuada" (*La Imagen Científica*, p. 28). Dando a entender que no hay ningún compromiso con la verdad de las teorías, y por lo tanto cualquier inferencia que hagamos tendiente a justificar algo más allá de lo empíricamente adecuado, es ilegítima. En nuestro caso, sólo tenemos una buena simulación... y nada más.

Sin embargo, con leer unos pocos trabajos de investigaciones recientes donde se utilizan modelos y simulaciones computacionales de supuestos procesos mentales¹, resultaría fácil advertir que se utiliza cierto criterio no explicitado que permite discriminar entre simulaciones más o menos reales. A ese principio se lo podría enunciar como: "sólo aquellas simulaciones que posean la capacidad de ser ejecutadas por una red neuronal (como la de un cerebro) son susceptibles de considerarse reales, todas las restantes caen en la categoría de meros artefactos computacionales".

Al elaborarse una simulación computacional, se involucran teorías². Y el éxito de la simulación opera como el certificado de la adecuación empírica de estas teorías. A esto deberemos sumarle, según el principio expuesto arriba, una segunda adecuación ya no empírica sino teórica: la simulación debe adecuarse al cuerpo de conocimientos mediante los cuales se describe el comportamiento de las redes neuronales existentes en el cerebro.

Habría entonces una versión más realista que otra de las simulaciones. Una donde tenemos más razones para creer en su realidad.

A simple vista parece haber triunfado el realismo, al fin y al cabo contamos con un criterio capaz de discriminar entre las simulaciones "reales" y las que no lo son. Pero la posición defendida por van Fraassen no termina aquí.

Para él, el hecho de que una de nuestras simulaciones sea considerada más real que otra carece totalmente de sustento. Lo único que logramos fue dar con la simulación que mejor se adapta al conjunto de creencias aceptadas como válidas en determinado momento. No existe nada por detrás de las teorías utilizadas que las haga ser más o menos exitosas, y por tanto no disminuimos un ápice la distancia a realidad alguna. Según esta posición, las teorías sólo deben salvar las apariencias de lo observable a "ojo desnudo", todo lo que ellas contengan como

suposiciones acerca del mundo real son puras herramientas que permiten *construir* el mundo *empírico* que observamos.

En este sentido, todo el cuerpo de conocimientos neurofisiológicos puede ser tomado como una creencia válida más (para determinado momento). Y el principio antes expuesto queda reducido a: "sólo aquellas simulaciones que posean la capacidad de adecuarse a estas creencias son susceptibles de considerarse reales...", donde la frase "considerarse reales" ha perdido toda significación salvo como explicitación de los caprichos científicos de una época particular, en la cual la realidad se encuentra construida y determinada por un conjunto de creencias.

van Fraassen nos deja en un callejón sin salida. Sólo es aparentemente real aquello que creemos que lo es en función de las cosas observadas a través de nuestros sentidos desnudos, todo lo que vaya más allá de ellos (observar por medio de un microscopio o un telescopio, por ejemplo) formará parte del mundo empírico construido por nosotros. Las teorías se *adaptan* mejor o peor a estos mundos construidos. Y eso es todo.

Esta manera de argumentar³ invalida la pregunta misma acerca de qué hace más o menos exitosa a una teoría como explicación de cierto fenómeno pues lo único que hay es un buen encastre entre la teoría y el "ambiente" teórico del momento. Sin embargo, nada hay que nos obligue a detener nuestras preguntas donde nos indica van Fraassen, de la misma manera que nada nos impide buscar las razones que hacen a cierto rasgo más o menos ventajoso a la hora de evaluar su valor adaptativo. Y es justamente en este sentido que vuelven a adquirir significado las preguntas iniciales acerca de la realidad involucrada dentro de las simulaciones, porque estamos indagando acerca de cuáles son los motivos que nos hacen aceptar como "mejor adaptadas" a la realidad (entendida, si queremos, en términos de creencias) ciertas simulaciones por sobre otras.

Lo que estoy intentando dejar en claro es que la propuesta de van Fraassen se queda corta a la hora de dar cuenta de nuestra pregunta, inclusive para el caso en que aceptamos la realidad "empíricamente construida" defendida por él.

Aunque la realidad sea construida, queremos saber cuánta de esa realidad se encuentra contenida en una simulación de procesos mentales.

Nancy Cartwright propone que para mostrar el "lado" antirrealista de una simulación, nos fijemos en cuál es la definición del diccionario. La que ella cita, la segunda acepción del *Oxford English Dictionary* dice de un simulacro (la traducción es mía): "algo que meramente mantiene la forma o la apariencia de cierta cosa, sin poseer su sustancia o cualidades propias"⁴. Pienso que hasta aquí van Fraassen aplaudiría. Pero Cartwright agrega a continuación "el éxito del modelo (la simulación) depende de cómo y cuán precisamente éste pueda replicar lo que sucede" (op. cit., p. 153), donde "lo que sucede" está indicando una presencia imposible de digerir para un estómago antirrealista. Sin dudas, el término introduce alguna ambigüedad. Y esta ambigüedad lejos de ser un defecto, pertenece al corazón de la posición de Cartwright, que se diferencia del antirrealismo de van Fraassen, según ella misma lo expone, en dos puntos: a) acepta entidades teóricas no restringidas a lo observable y b) acepta leyes fenomenológicas que se refieren a causas "reales". La coincidencia entre ellos se encuentra en sus consideraciones

acerca de las idealizaciones que permiten a las teorías funcionar, que en el caso de Cartwright dichas idealizaciones abarcan tan sólo las leyes *fundamentales*.

Entonces, de resultar exitosa nuestra simulación, podríamos sostener que algunos de los conceptos involucrados en la construcción del simulacro (sólo las leyes causales), se refieren a la realidad, mientras que otros son sólo idealizaciones que no poseen ningún referente real.

Según Cartwright "en la explicación por simulación el modelo es la teoría del fenómeno" (op. cit., p. 159), y dicho modelo se relaciona por un lado con el mundo y por el otro con las herramientas de cálculo matemático. Por ejemplo, en la simulación de los procesos mentales habrá términos que se referirán a las leyes fenomenológicas causales que se *observan* en las experimentaciones con neuronas, pero bajo la forma *idealizada* que permite su tratamiento matemático. También encontraremos términos que "convienen" se encuentren en ese lugar y que en sí mismos no refieren a nada real: solamente se utilizan porque permiten el tratamiento computacional de la teoría.

Para Cartwright, entre el mundo y el modelo simulado se tienden *puentes*, puentes cuya realidad se evalúa según el criterio de causalidad, que es en última instancia el que determina también qué leyes son reales (leyes fenomenológicas) y cuales son meras generalizaciones organizativas o idealizaciones matemáticas (leyes fundamentales).

El problema manifiesto de esta perspectiva surge como resultado de la adopción de la causalidad como criterio que permite diferenciar lo real de lo ideal y en última instancia también lo observable de lo inobservable. ¿Es acaso la causalidad ella misma observable? van Fraassen se mantiene escéptico respecto a cualquier criterio que permita distinguir lo observable de lo que no lo es, pagando el precio de circunscribir la ciencia a las "apariencias" observables a simple vista, con lo que imposibilita cualquier argumento que pretenda explicar por qué una teoría es más real que otra.

Conclusión

De una manera muy resumida, he intentado aplicar los conceptos provenientes de la discusión entre realistas y antirrealistas al campo de las simulaciones computacionales de los procesos mentales, encontrando que el antirrealismo de Nancy Cartwright, menos radical que el de van Fraassen, permite relacionar el éxito de una simulación con ciertas características de la realidad, restringidas por cierto a las observaciones causales que podamos efectuar en nuestros laboratorios. Lo cual indicaría de alguna manera que sólo en la medida en que desarrollemos estas observaciones podremos tender los puentes entre nuestras simulaciones computacionales y la realidad del funcionamiento de nuestra mente.

Sin embargo, aún se escucha un eco recordándonos que eso que hacemos en el laboratorio puede ser entendido como la construcción del mundo empírico al cual ajustaremos nuestras simulaciones. El resultado global es que nuestras teorías se adaptan a eso que llamamos realidad y eso que llamamos realidad se construye por medio de nuestras teorías. Así, el campo de las simulaciones computacionales aparece idéntico a cualquier otra forma de producir conocimiento. Al menos respecto a sus relaciones con la realidad.

En cuanto a la pregunta ¿cuánta realidad hay en una simulación de procesos mentales?, deberíamos responder: depende. Depende de qué consideremos en cada momento como real.

Notas

¹ En una vasta bibliografía consultada para un trabajo sobre percepción de la constancia del movimiento (más de treinta artículos, principalmente de Vision Research y de Perception), encontré que todos los autores intercalaban en algún punto de su argumentación referencias a estudios neurofisiológicos estimados como coherentes con sus hipótesis. En la bibliografía se exponen algunos de estos trabajos.

² Por ejemplo, si queremos simular el movimiento de nuestro sistema solar, echaremos mano a las leyes de Newton para calcular las trayectorias que los cuerpos describirán en la pantalla (y si queremos mayor precisión, a las leyes de Einstein). También necesitaremos conocimientos de geometría para respetar las escalas a las que deben aparecer los planetas, etcétera.

³ Me refiero al recurso adaptacionista o, mejor dicho, evolucionista que utiliza van Fraassen.

⁴ Quizás sea el lugar adecuado para hacer notar un aspecto interesante de las simulaciones. En cuanto simulaciones computacionales, llevan implícita la característica de ser digitales. Y el hecho de que existan convertidores analógico/digitales y viceversa (el de los monitores CRT -DAC- y el de los scanners son ejemplos de ellos), nos habla de que hay al menos otro tipo de simulación: una de características analógicas. La mejor manera de comprender la diferencia entre una y otra es apelando a los conceptos de lo continuo y lo discreto. Nada mejor que un ejemplo.

Se dice comúnmente que un reloj analógico es aquel que representa "el camino del sol alrededor de la tierra" (aunque sabemos que es al revés). Así, los clásicos relojes de aguja son instancias de esta representación analógica: las agujas giran y representan el movimiento percibido del sol. Sin embargo, es necesaria la partición *digital* del cuadrante para poder efectuar la lectura de la hora: hace falta realizar marcas que representen los segundos, minutos y horas. El reloj sin marcas constituye una simulación analógica del paso del tiempo. Los relojes que sólo nos indican la hora mediante números son simulaciones digitales del paso del tiempo.

Adviértase que gracias a que el cuadrante del reloj analógico fue marcado, es posible la conversión de un sistema en el otro.

El interés de todo esto es llamarnos la atención acerca de la dificultad de simular cualquier proceso mediante un procedimiento digital: dicha dificultad es idéntica a la de representar lo continuo mediante lo discreto.

Bibliografía

(Específica sobre realismo / antirrealismo)

Bas van Fraassen, *La Imagen Científica*, México, UNAM, 1996.

Nancy Cartwright, *How the Laws of Physics Lie*, Nueva York, Oxford University Press, 1989.

A. Fine, "The Natural Ontological Attitude". En: *Readings in the philosophy of science - From positivism to postmodernism*, Th. Schick (comp.), London - Toronto: Mayfield Publ. Co, 2000.

(Algunos artículos consultados donde aparecen referencias a simulaciones relacionadas con conocimientos neurofisiológicos)

Barraza José F., Grzywacz Norberto M. "Measurement of angular velocity in the perception of rotation", *Vision Research* 42, pp 2457 - 2462, 2002.

Zohary Ehud, Sittig C. Anne. "Mechanisms of velocity constancy", *Vision Research* Vol 33 N°17, pp. 2467 - 2578, 1993.

Distler K. Hartwig, Gegenfurtner Karl R., van Veen Hendrik A.H.C., Hawken Michael J. "Velocity constancy in a virtual environment", *Perception* volume 29, pp. 1423-1435, 2000.