

CASTRO CALDAS, Alexandre. *Inteligência Vital, Estupidez Artificial*. Prefacio de Álvaro Laborinho Lúcio. Lisboa: Contraponto, 2025, 187 pp. ISBN: 9789896665050.

Si se quiere entender de dónde proviene la creencia –hoy generalizada y cada vez más popular– en una conciencia artificial, el reciente libro de Alexandre Castro Caldas plantea un buen punto de partida. Estamos ante un recorrido histórico sobre cómo ha ido evolucionando la noción de inteligencia. A veces, se presenta como una mirada crítica sobre los distintos análisis científicos que se han sucedido desde la Antigüedad; otras veces, como una simple muestra de lo que ha desarrollado en el campo de la investigación neurocientífica.

El volumen se compone de trece capítulos que orientan el tema a través de las preocupaciones del autor, principalmente sobre la definición de *inteligencia*; sobre si sería posible medir, de algún modo, ese fenómeno; la memoria y la imaginación, entre otras, bajo títulos que pueden ser considerados problemáticos. Sirva de ejemplo el tercer capítulo, titulado *A Redescoberta do Homem* (El Redescubrimiento del Hombre) (61-83), para referirse al ser humano en general. Francis Galton es destacado por Castro Caldas como autor de esquemas científicos por estirpes, en los que se idealiza una rama específica –y a la vez universal– de personas para analizar problemas del ámbito intelectual. Pero el problema no se agota en Galton. De hecho, se trata de un problema fundamentalmente moderno. Primero con René Descartes, seguido por Isaac Newton, y agravado años después por Charles Darwin. Es

decir, la cristalización de un modelo de sujeto, junto con la noción de un mundo entendido como *naturaleza*. El mundo se convierte en objeto-objetivo, en relación con un sujeto que no lo determina de ningún modo, solo lo percibe de forma aislada.

Nos encontramos con una pretensión transversal a lo largo de toda la obra: la de que existen datos ya resueltos –como la idea de fuerza o masa–, aunque no sea así, incurriendo en contradicciones. Paso a citar un ejemplo muy claro:

La inteligencia debe ser la capacidad de percibir el mundo y sus reglas, y de a ellas adaptarse biológicamente, en el sentido de mejorar la conservación individual y global; la conciencia es comprender cómo eso es hecho y, a partir de ahí, conocer el mundo y sus reglas (no todas) y sobre ellas dialogar internamente, utilizando metáforas (símbolos) o pantomima (gestos) y, a través de esos códigos, relacionarse con los otros (28).

El problema está en la reducción de la conciencia a una mera toma-de-conciencia, en el sentido popular de la expresión. Se tiene esta, aún, como un modo de análisis que se remite al mundo, posteriormente al hecho de este. Percibir-el-mundo, al contrario de lo que Castro Caldas y otros afirman, sería mejor visto como un proceso existencial que, de ningún modo, consigue ser formulado en tanto proceso no-significante del propio Mundo. Es decir, que lo consciente es constituyente del mundo, en la misma medida en que el mundo es constituyente de la conciencia. «Dialo-

gar internamente», como escrito por el autor, se presenta como una débil descripción del proceso significativo consciente. ¿Sería el diálogo algo posible de ser reducido a una idiosincrasia aislada, de modo interno, sin ningún tipo de expresividad (externa)?

Siendo los primeros capítulos una propuesta de una breve revisión histórica del concepto de *mente*, es incomprensible hacer parecer que solo el campo de ciencias como la biología, la psicología o, más recientemente, la neurociencia, se muestra valioso –aunque haga breves referencias muy superficiales a Aristóteles o Platón–. Esta gran carencia por parte del autor viene de una aparente falta de familiaridad con siglos de literatura filosófica, especialmente fenomenológica, dando con Hegel los primeros signos, ganando enorme fuerza con Husserl y la tradición posterior de Heidegger y Sartre, predominantemente. En todo caso, los grandes avances, fundamentales para el pensamiento contemporáneo, junto con los principales autores de la historia de la filosofía, son como dejados en segundo plano, brevemente referenciados o no referenciados en absoluto.

Considerando que una y cualquier actividad observadora *está-con* el ser humano, nunca podremos olvidar que este es aquel quien genera inferencias sobre determinada cosa –o estado de cosas– con la cual se está relacionando en determinado momento. Esto incluye, por supuesto, durante la actividad científica. En otras palabras, el análisis científico solo tiene sentido, fundamentalmente, por dependencia de una base antecedente instigadora de estructuras significantes [el ser humano]. Merleau-Ponty dejó este hecho bastante claro. Castro Caldas

aparenta saberlo, pero aunque hace una breve referencia a Merleau-Ponty en las primeras páginas, acaba por seguir la mayor parte del libro sin aparente noción de lo que se puede aprender con los textos del filósofo. Es curioso, además, que valore las teorías freudianas después de publicadas las oposiciones de Jean-Paul Sartre al inconsciente freudiano, ignorándolas completamente. Vimos con Sartre que, para un análisis de la conciencia, las teorías freudianas se revelan como inutilizables, no solo por la equivocada compartimentación de la posibilidad subjetiva [consciente | inconsciente], sino también por la cristalización en un *ego* fijo de aquello que siempre se renueva. Exponer debidamente los argumentos de Sartre sería asunto para otro espacio. Mi referencia al autor francés viene a propósito de no entenderse por qué se excluyen autores tan relevantes para un cuestionamiento sobre la conciencia, la *inteligencia* y la *estupidez*. Muchos de ellos con argumentos convincentes contra varios de los autores presentados. Esta sería la principal carencia de *Inteligência Vital, Estupidez Artificial*.

Es posible señalar aún otro gran problema, esta vez en la segunda parte del libro, a partir del quinto capítulo. Importa evidenciar que Castro Caldas se refiere al concepto de teoría de la mente, solo como una capacidad de «tener conciencia de los propios estados mentales y de los de otras personas» (100). Indica que la noción fue propuesta por primera vez por Premack y Woodruff, en 1978, con respecto a los chimpancés, cuando se descubrió que estos eran capaces de proyectar posibilidades que los trascendían, como comprender que al

conectar el cable de un aparato eléctrico a la corriente, este funcionaría (99). Esta noción sería posteriormente desarrollada en detalle por Henry Wellman (psicólogo), quien publicó una profunda revisión de sus investigaciones en 2018. En realidad, la discusión sobre la hipótesis de una teoría de la mente se divide más propiamente entre 1. «theory theory of mind» [TT] y 2. «simulation theory of mind» [ST], que no son exactamente lo mismo, habiendo, incluso, autores de teorías híbridas de estas. Una aclaración por parte de Castro Caldas acerca de en qué lado se posiciona habría sido de gran utilidad. Por el ejemplo presentado —la conocida *prueba de la falsa-creencia*—, parece estar refiriéndose a la primera forma de una teoría de la mente [TT]. Explicándola brevemente: Se le pide a un niño que prevea si otro (a) buscará un objeto donde previamente creía que estaba guardado, o (b) si irá a buscar ese mismo objeto al lugar donde un tercero lo escondió y cambió de sitio (100). Se intenta comprender, a través de esta prueba, si el niño concibe al otro como distinto en sus creencias, mediante una teorización de los respectivos estados mentales, constituyéndose así como una subjetividad única. La TT propone, esencialmente, una comprensión del mundo como teóricamente mediada, por etapas. Se diría, instrumental. Lo que significa que, descartando dicho instrumento —capacidad de teorización—, una comprensión de la propia experiencia, así como de la del otro, se vuelve imposible. Se asume, a través de este modelo, que un niño no logra constituir ninguna noción subjetiva durante los tres primeros años de vida. Nos enfrentamos aquí a una dependencia de

procesos intelectuales para reconocer una *individualidad* en el mundo, donde el cuerpo queda ausente en su importancia para la constitución de la primera y, en conjunto, de lo último. Los estados mentales serían, en este aspecto, inobservables —ocultos a los sentidos (Gallagher y Zahavi 2008, 172)— alcanzados solo si ponemos en marcha un esfuerzo teórico. El cuerpo deja de ser relevante como posibilidad expresiva y directa de esos estados «mentales», ya que la decodificación del mundo en el que uno se encuentra, co-habitado por otros, reside en el «interior» de uno mismo —radicalizando una posición distanciada (cf. Varela *et al.* 1993, §IV). Se trata de un «instrumento cognitivo» (101), como escribe Castro Caldas.

Paso a evidenciar el segundo problema de forma más clara: como la inteligencia, junto con la conciencia del individuo, es propuesta por el autor como, esencialmente, un simple receptor o decodificador del Mundo exterior —en lugar de un constante proceso afectivo que compromete al ser humano—, también la estupidez termina por convertirse en un efecto que surge desde dentro del ser humano, como «resultado de una cierta forma de funcionamiento cerebral» (134). Contrariamente, yo argumentaría que la estupidez no es más que una inadecuación del individuo en relación con una práctica común. Eso justificaría algo con lo que Castro Caldas mínimamente estaría de acuerdo, cuando dice que la estupidez «resulta de la dificultad de interpretar las informaciones del mundo generando por ello creencias falsas» (139). Hay una referencia a una relación del individuo con su exterior. Sin embargo, Castro Caldas considera

que el fallo *estúpido* ocurre meramente por falta de analogía con una experiencia pasada. Se excluye nuevamente la importancia del cuerpo y de la práctica social, siempre actualizables, para consideraciones relativas a procesos de conciencia, muchas veces no explicables si recurrimos solamente a teorías de la cognición. Cualquiera, que viaje a un país con una cultura radicalmente distinta a la propia, se encuentra sujeto a ser tratado como estúpido en sus acciones, pues actuará según una práctica social completamente diferente, donde una experiencia de vida, junto con la memoria de esta, no son suficientes para salvarse de la humillación.

Para concluir, vale la pena revisar los últimos tres capítulos (153-174). En ellos, el autor presenta la *Inteligencia Artificial* a los lectores, planteando varios problemas que surgen como consecuencia de su realidad, pero sin reflexionar de manera muy crítica sobre ella. En ciertos momentos, subraya la importancia del *gut feeling* por parte del médico (166). Se olvida, no obstante, de que ignorar ese *gut feeling* como algo importante para el desarrollo social y humano, protegiendo su vitalidad, es una consecuencia natural de una visión tecnicista sobre la mente humana: si reducimos la mente, aunque dependiente del cuerpo, a una serie de impulsos eléctricos y variaciones químicas que ocurren en el cerebro, entonces producir una máquina que soporte esas transmisiones, con un análisis perfecto de un mundo de naturaleza y realidad objetivas, es una consecuencia comprensible y, en cierto punto, deseable. Castro Caldas comprende que existen problemas con esta posibilidad, pero entender el origen de esos problemas se vuelve

inalcanzable solo a partir del libro que presenta. Quizás porque el origen está en la misma corriente de pensamiento que apoya en varios momentos.

Las observaciones presentes en el último capítulo (173-174) reflejan la esencia del libro en su conjunto: una serie de consideraciones sobre temas de extrema profundidad filosófica, ignorándolos como tal, ofreciendo respuestas poco convincentes y, en varios casos, problemáticas. Es una lectura que presenta una revisión histórica de nociones fundamentales para la comprensión contemporánea de fenómenos conscientes, sin contrastarlas con las principales corrientes filosóficas que han surgido a lo largo de la historia. *Inteligencia Vital, Estupidez Artificial*, de Alexandre Castro Caldas, es un resultado de una tradición positivista y fisicalista, que en cierto punto considera el cuerpo como fundamental del proceso consciente, pero que termina por considerarlo, aún, en tanto que un aparato fragmentado.

REFERENCIAS BIBLIOGRÁFICAS

- GALLAGHER, Shaun, y Dan ZAHAVI. *The Phenomenological Mind. An Introduction to Philosophy of Mind and Cognitive Science*. Nueva York: Routledge, 2008.
- VARELA, Francisco, Evan THOMPSON, y Eleanor ROSCH. *The Embodied Mind: Cognitive Science and Human Experience*. Massachusetts: MIT Press, 1993.

Daniel CASTRO LOBO
 Universidad de Salamanca
 ORCID: <https://orcid.org/0009-0007-8408-8554>