



El volcado de la mente en la máquina y el problema de la identidad personal

Mind uploading and personal identity

 **Antonio Diéguez**
dieguez@uma.es
Universidad de Málaga, España

Recepción: 31 Mayo 2022
Aprobación: 29 Julio 2022
Publicación: 01 Diciembre 2022

Cita sugerida: Diéguez, A. (2022). El volcado de la mente en la máquina y el problema de la identidad personal. *Revista de Filosofía (La Plata)*, 52(2), e054. <https://doi.org/10.24215/29533392e054>

Resumen: En este trabajo se analiza la cuestión de si el volcado de la mente en una máquina (*mind uploading*), en caso de ser alguna vez tecnológicamente posible, mantendría o destruiría la identidad personal de quien experimentara el volcado. Se verá cómo podría contestarse a la cuestión en función de los criterios para el mantenimiento de la identidad personal que se asuman. No hay una respuesta única, puesto que la identidad personal se mantendría o no en función de los supuestos aceptados. No obstante, para la mayor parte de las concepciones filosóficas sobre la identidad personal, ésta no se mantendría en un proceso de volcado de la mente en una máquina. Si esto es así, habría que concluir que, desde diversos enfoques, dicho volcado sería equivalente a la desaparición del sujeto, es decir, a su muerte, y solo desde posiciones muy limitadas podría defenderse lo contrario.

Palabras clave: Volcado de la mente, Identidad personal, Transhumanismo, Inmortalidad tecnológica, Mejoramiento humano.

Abstract: This paper explores the question of whether mind uploading, if ever technologically possible, would maintain or destroy the personal identity of whoever experienced the uploading. It will be explained how the question could be answered in different ways depending on the criteria for the maintenance of personal identity that are assumed. There is no single answer, since personal identity would be considered as maintained or not depending on the accepted assumptions. However, for most of the philosophical conceptions about personal identity, this would not be maintained in a process of uploading. This means, therefore, that this process would be equivalent to the disappearance of the subject, that is, to his death, and only from very limited positions could the opposite be defended.

Keywords: Mind Uploading, Personal Identity, Transhumanism, Technological Immortality, Human Enhancement.

0. Introducción

En este trabajo analizaremos la cuestión de si el volcado de la mente en una máquina, en caso de que sea alguna vez tecnológicamente posible, mantendría o destruiría la identidad personal de quien experimentara el volcado. Este tema es un tópico habitual en la literatura filosófica acerca de las promesas del



EDICIONES
DE LA FAHCE



transhumanismo, que permite profundizar desde una nueva perspectiva en la cuestión filosófica de la identidad personal. De hecho, se trata de un tema de gran interés para el análisis de esta cuestión, puesto que como bien ha subrayado Patrick D. Hopkins (2021, p. 233):

El problema con el volcado, en esencia, es solo una versión actualizada marginal del viejo problema de la identidad personal [...]. La cuestión real del volcado es si los procedimientos para conseguirlo mantienen la identidad de la mente específica a lo largo del proceso.

Aquí analizaremos cómo podría contestarse a la cuestión en función de los criterios para el mantenimiento de la identidad personal que se asuman. Veremos que no hay una respuesta única, puesto la identidad personal se mantendría o no según las implicaciones de los supuestos que se aceptaran. No obstante, puede decirse que, para la mayor parte de las concepciones filosóficas sobre la identidad personal, ésta no se mantendría en un proceso de volcado de la mente en una máquina. Esto significa que, de acuerdo con ellas, dicho volcado sería equivalente a la desaparición del sujeto, es decir, a su muerte.

1. Las dificultades del volcado de la mente

Entre los objetivos a medio o largo plazo del transhumanismo está la unión progresiva del ser humano con las máquinas, siendo la fase final de esta unión, al menos para los más radicales, cuyo anhelo es distanciarse lo más posible del ser humano actual, el volcado de la mente en un ordenador lo suficientemente complejo (*mind uploading*). Se ha dicho, como objeción a este proyecto, que incluso si fuera técnicamente posible el volcado (cosa bastante dudosa por el momento), o filosóficamente posible (cosa también dudosa, puesto que para buena parte de la filosofía actual la mente no es un software que pueda funcionar igual con independencia del soporte material), la consciencia y la identidad personal del sujeto implicado no tendría por qué mantenerse tras esa fusión. Para los críticos, por tanto, la mente no solo no sería algo descargable, sino que en caso de que, *per impossibile*, la tecnología lo permitiera alguna vez, tal cambio no serviría como medio de supervivencia de los sujetos, sino que los sujetos desaparecerían en el proceso (Häggstrom, 2021).

Conviene hacer en este punto una precisión. Hay que distinguir la emulación completa del cerebro (*Whole Brain Emulation* (WBE)), que es, por ejemplo, lo que se está intentando hacer con el nematodo *Caenorhabditis elegans*, del volcado de la mente en una máquina. Un volcado genuino y no fallido no podría limitarse a emular el funcionamiento de un cerebro, sino que exigiría mantenimiento de la consciencia y de la identidad personal (Chalmers, 2014). Y conviene saber también que ni siquiera en el caso relativamente simple de *C. elegans*, con un sistema nervioso de 302 neuronas, se ha conseguido, pese a lo que a veces se dice, una simulación completa de dicho sistema nervioso, no digamos ya un volcado de su “mente” (Niconiconi, 2021). Con el ser humano las cosas serían mucho más complicadas. El cerebro humano contiene 86 mil millones de neuronas y se estima que el número de sus conexiones sinápticas es de $1,5 \times 10^{14}$ (un 15 seguido de 14 ceros), siendo estas conexiones, además, cambiantes. Añádase a eso que, como hoy bien se sabe, el conectoma no lo es todo a la hora de

configurar nuestros procesos mentales; son muy importantes también las células gliales, sin las cuales el cerebro no podría funcionar. En los seres humanos, según estimaciones recientes, casi igualan el número de neuronas (Azevedo et al., 2009, von Bartheld, 2018). ¿Qué capacidad de computación y almacenamiento tendría que tener un ordenador para procesar toda esta información de un solo cerebro? Seguramente una cifra mareante de terabytes, muy por encima en todo caso de la capacidad de cualquier ordenador previsible con la tecnología actual.

De forma bastante convincente, Jesús Zamora Bonilla ha argumentado que si estas dificultades hacen ya enormemente complicado desde un punto de vista tecnológico el copiado de una mente (de un cerebro) en todos sus detalles, mucho más difícil sería conseguir que esa mente copiada funcionara en realidad como la mente original. Téngase en cuenta que un pequeño porcentaje de error en la descarga, daría lugar a algo funcionalmente muy distinto del cerebro original. Zamora Bonilla escribe:

Si la posibilidad de escanear un cerebro vivo con todo el detalle necesario parece remotísima, la de fabricar un cerebro que fuese una réplica neurona por neurona, sinapsis por sinapsis, de otro cerebro que hubiéramos escaneado previamente creo que está fuera de los límites de cualquier tecnología posible. (Zamora Bonilla, 2019, p. 22).

Para que la mente individual pudiera ser reactivada tal y como ésta es no bastaría con copiar toda la estructura o información que la constituye, sino que habría que copiar también su modo de interacción con el organismo y con el entorno, o, como Zamora Bonilla dice recurriendo a Ortega, habría que reconstruir también “su circunstancia”. La razón es simple:

Las capacidades y estados mentales de cualquier animal, incluidos los seres humanos, no dependen solo de lo que sucede en el cerebro, sino también de las peculiaridades de nuestros organismos y de los seres que nos rodean, en interacción con los cuales han evolucionado durante millones de años. La mente es «corpórea», «extendida». (Zamora Bonilla, 2019, p. 22).

Esta idea de que nuestra mente es siempre una mente corporizada e incluso “extendida” en elementos que nos rodean y que están fuera de nuestro cerebro, incluyendo, por supuesto, los recursos tecnológicos que tenemos hoy a nuestra disposición, goza de una creciente aceptación. Desde esta perspectiva, mi mente no “está” solo en mi cerebro; mi mente incluye también mi cuerpo y algunas de las cosas de mi entorno con las que interacciono cognitivamente. Las interacciones sensomotoras con mi entorno forman parte central de ella y de mi identidad como agente. Todo eso cambiaría de forma radical después de un volcado (Clark y Chalmers, 2011).

Un problema adicional es el que plantea la naturaleza cambiante del yo, que también debería ser preservada. El yo no es pasivo ni inmutable, sino que está en permanente autoconstrucción (DeGrazia, 2005). Ortega ya lo decía de una forma más sonora: el ser humano ha de ser el novelista de sí mismo. Cada uno de nosotros tiene que hacerse a sí mismo, en una tarea indelegable, en buena medida narrativa, a partir de un proyecto vital que debe crear de forma personal. No nacemos con un yo, con una identidad, acabada, completa. Más bien al contrario, nuestra vida consiste en la tensión inacabable de lograr completar esa construcción del yo que debemos ser en una (o contra una) circunstancia dada,

imaginándolo de forma creativa o –si somos incapaces de eso– copiándolo a otros, y resignándonos entonces a una vida inauténtica, que es tanto como decir a una no-vida. De ahí que vivamos siempre orientados hacia el futuro, aunque en proximidad con el pasado. En su obra *Goya*, Ortega lo explica del siguiente modo:

El yo es, pues, lo más irrevocable en nosotros. Pero esto no implica que no varíe. Nuestro yo no es, por fuerza, siempre idéntico. Al contrario, experimenta mutaciones que a veces son radicales y tampoco provienen de nuestro albedrío, sino que se producen en él, más o menos motivadas por experiencias de la vida, pero, en última instancia, con un carácter de inexorable espontaneidad. (Ortega y Gasset, 1964, p. 549).

La cuestión es si mi mente volcada en una máquina sería capaz de seguir llevando adelante esa tarea de autoconstrucción del yo. No es algo que deba darse tampoco por sentado. E incluso si pudiera continuar con ella, no está claro que no hubiera una ruptura radical con el proyecto de vida anterior. Más bien, es probable que lo hubiera.

2. El volcado de la mente: ¿en qué consiste?

Mark Walker (2014) ha resumido muy bien las tesis con las que ha de comprometerse todo aquel que acepte la posibilidad del volcado de la mente en una máquina. Serían las al menos las siguientes:

1. Las computadoras son capaces de soportar las importantes propiedades constitutivas de la identidad personal, por ejemplo, el pensamiento y la consciencia.
2. Es posible capturar la información necesaria para emular las propiedades importantes de los humanos individuales.
3. Es posible sobrevivir al proceso de carga.

Supongamos, por mor del argumento, que alguna vez la tecnología estuviese lo suficientemente desarrollada como para hacer posible el volcado de la mente, es decir, como para salvar los obstáculos planteados por (1) y (2) y conseguir un duplicado equivalente funcionalmente al cerebro de una persona que tuviera consciencia y pleno uso de sus facultades mentales. La cuestión que vamos a considerar aquí es la tercera, que, en última instancia, sería justamente la cuestión del mantenimiento de la identidad personal después del volcado.

Ante todo, hay que admitir que no todos los que han tratado esta cuestión estarían de acuerdo en que la supervivencia del individuo implica necesariamente el mantenimiento de la identidad personal. Es sabido que Derek Parfit lo niega. Para él, la supervivencia (tras el volcado) sería lo realmente importante y el mantenimiento de la identidad personal solo algo secundario, algo, en el fondo, meramente conceptual, que no tendría por qué darse y de lo que podría prescindirse. Lo que importa es que haya continuidad psicológica entre el sujeto y la mente volcada, y no si al resultado puede seguir siendo considerado como esa misma persona. Estas son sus palabras:

Desde su punto de vista [el de mucha gente], lo que nos da una razón para preocuparnos por nuestro futuro es, precisamente, que será nuestro futuro. La identidad personal es lo que importa en la supervivencia.

Yo rechazo esta postura. Creo que la mayor parte de lo que importa son otras dos relaciones: la continuidad y la conectividad psicológicas que, en los casos ordinarios, se mantienen entre las diferentes partes de la vida de una persona. Estas relaciones solo coinciden aproximadamente con la identidad personal, ya que, al contrario que la identidad, son, en parte, cuestiones de grado. Y creo que tampoco importan tanto como se supone que lo hace la identidad. (Parfit, 2004a, p. 108).

Parfit ilustra esta tesis en principio sorprendente con el experimento mental de la mente escindida. Supongamos que cada uno de nuestros hemisferios cerebrales por separado pudiera (quizás mediante cierta ayuda tecnológica) hacerse cargo de todo nuestro cuerpo y de nuestra mente, algo que no parece descabellado si tenemos en cuenta que hay personas que viven con bastante normalidad con un solo hemisferio cerebral funcional. Supongamos que padecemos una enfermedad mortal y se nos opera para extraer cada uno de nuestros hemisferios por separado con objeto de trasplantarlos a dos cuerpos diferentes, y que finalmente ambos cuerpos sobreviven y funcionan bien con esos hemisferios cerebrales ¿cuál de los dos nuevos individuos soy yo realmente? No tendría ninguna buena razón para decir que solo soy uno de ellos en lugar del otro. Pero tampoco podría decir que los dos sean yo, porque esto sería absurdo: no puede haber dos personas que puedan decir con propiedad que son yo. Por lo tanto, este caso mostraría que, suponiendo que no quiera simplemente desaparecer, lo que importaría en última instancia es que he sobrevivido en dos cuerpos distintos, y no sería tan relevante si mi identidad personal no hubiera quedado preservada tal como era.

La idea es interesante y el argumento que ofrece Parfit tiene cierta fuerza, sin embargo, no sería exagerado decir que no ha convencido a muchos (Kind, 2015, pp. 62-65). Para la mayor parte de los analistas parece claro que si lo que “sobrevive” después del volcado no soy yo, en algún sentido fuerte (y no meramente como una cuestión conceptual), el volcado sería equivalente a mi muerte. Y no son pocos los que también creen que, si bien la continuidad psicológica de la que habla Parfit constituye una condición necesaria para el mantenimiento de la identidad personal, no es, sin embargo, una condición suficiente. No basta, como se ve en el ejemplo de la mente escindida, con que la continuidad psicológica no se rompa. Se han de cumplir requisitos adicionales (como el de la no-división). Para algunos críticos, esto llevaría a concluir que, en el caso de la mente escindida, el sujeto como tal ha desaparecido después de la división y lo que ha surgido en cada cuerpo son dos nuevos individuos distintos.

Un problema fundamental a la hora de dilucidar esta cuestión de forma provechosa sigue teniendo difícil respuesta: no solo no está claro cómo sería una fusión futura con las máquinas, sino que tampoco hay acuerdo acerca de en qué consiste la identidad personal y cuáles son los criterios de su permanencia en el tiempo. Por ello, a continuación, vamos a analizar las diferentes posibilidades de permanencia de la identidad personal después de un eventual volcado de la mente dependiendo de las diferentes concepciones vigentes sobre el mantenimiento de la identidad personal.

Hablar del volcado de la mente significa realizar la transferencia de la mente (pero, ¿de qué en concreto?) de un cerebro humano a una máquina, o bien de una máquina que ya tenga una mente a otra máquina (Häggström, 2021). Algo así implicaría, como hemos señalado, la copia exacta de nuestro cerebro, de todo su conectoma, pero también de los detalles concretos de cada sinapsis, y la puesta en funcionamiento de esa copia mediante los procedimientos oportunos (y esto

abriría la posibilidad de hacer más de una copia de nuestra mente, al menos en principio). Se ha hecho popular entre los interesados en el tema la distinción de David Chalmers (2010 y 2014) entre volcado destructivo, volcado gradual y volcado no destructivo, aunque es una distinción extraña, porque, en realidad, el volcado gradual podría caer bajo cualquiera de las otras dos categorías.

El volcado destructivo implica la destrucción en el mismo momento del volcado del cerebro original cuya mente ha sido volcada. El volcado gradual, que se conseguiría supuestamente mediante nanotecnología, sería la sustitución una a una de todas nuestras neuronas por neuronas artificiales que, con anterioridad, habrían copiado el funcionamiento exacto de la neurona a la que sustituyen, o bien podría consistir también en el volcado funcional de cada neurona en una máquina conectada al cerebro, de modo que la máquina asumiera la función de esas neuronas volcadas y pudiera sustituirlas. En cierto modo, es una modalidad del volcado destructivo, puesto que se supone que la neurona sustituida es destruida, pero cabría la posibilidad (por remota que fuera) de que se conservara la neurona sustituida y fuera unida a todas las neuronas que se fueran sustituyendo gradualmente para así recomponer finalmente el cerebro original, o también cabría la posibilidad de que se copiara el cerebro neurona a neurona, pero dejándolo intacto en el proceso. El volcado no-destructivo consistiría en la copia (gradual o simultánea) de toda la estructura cerebral y la información contenida en ella sin destruir el cerebro original. Chalmers cree que la versión más factible sería la del volcado gradual. Aunque el tema es complejo y tiene ciertamente interés filosófico, no entraremos aquí en su discusión. Aceptaremos la tesis de Chalmers de que la alternativa más favorable para el mantenimiento de la identidad personal sería el volcado gradual (destructivo), si bien, como él mismo señala, un volcado gradual que pudiera hacerse en un tiempo muy breve (minutos o segundos) diferiría probablemente muy poco de un volcado instantáneo.

Chalmers piensa que en el proceso de volcado gradual se mantendría la consciencia, o al menos cree que esa es “de lejos la hipótesis más plausible”, puesto que la consciencia es un “invariante organizacional, es decir, que los sistemas con el mismo patrón de organización causal tienen los mismos estados de consciencia, sin que importe que la organización esté implementada en neuronas, silicio o algún otro sustrato” (Chalmers 2014, p. 107). Esto no es más que una redescipción de la bien conocida (y controvertida) tesis funcionalista en filosofía de la mente, defendida por Hilary Putnam y Jerry Fodor, entre otros. En definitiva, esta tesis ve la mente como una estructura informacional de inputs y outputs, como un software, en suma, que puede ejecutarse en hardwares diferentes sin experimentar cambios en su estructura. De forma más precisa, los procesos mentales son procesos internos que median causalmente entre las entradas y salidas del sistema y se caracterizan por su función, con independencia del soporte material del sistema. No obstante, también desde el dualismo mente/cuerpo podría mantenerse algo similar a lo que sostiene Chalmers.

Ahora bien, incluso aunque se aceptara que la consciencia es un invariante organizacional y puede ser preservada, por tanto, en el volcado a una máquina, quedaría por resolver la cuestión de si el resultado del volcado, aunque fuera consciente de sí mismo, seguiría siendo también la misma persona, teniendo en cuenta que la identidad personal, como admite igualmente Chalmers, no es un invariante organizacional. Una copia de mi mente podría ser cualitativa

y organizacionalmente idéntica a mi mente y, aun así, puesto que sería numéricamente distinta, no sería yo mismo. Por eso, como señala Patrick Hopkins, para admitir que el volcado de la mente pueda funcionar, habría que mostrar que el acto de volcar una mente es muy diferente en su naturaleza al acto de copiar un objeto, puesto que, si el volcado de mi mente es como realizar una copia de un objeto, no existiría entre mi mente y su copia una identidad en sentido estricto (Hopkins, 2012, p. 237).

Este es el problema central que ahora hemos de discutir aquí, el de la persistencia de la identidad personal. ¿Qué hace que una persona en el tiempo “t” sea la misma que una persona en el tiempo ‘t’? ¿Podríamos decir que la identidad personal se mantendría tras un volcado de la mente en una máquina? ¿Pensarán los demás que el resultado de ese volcado soy yo? ¿Tendría eso relevancia? ¿Pensará mi mente una vez volcada que soy yo? Y si hubiera varias copias ¿pensaran esas copias que son yo? ¿Podría yo pensar ahora en ellas como siendo todas yo?

Este problema debe ser distinguido del problema *estricto* de la identidad personal, que consistiría en establecer las condiciones que hacen que una persona sea ella misma, o si se quiere, las propiedades que individualizan a una persona (Macdonald, 2005, cap. 4). En este sentido, se ha distinguido entre la identidad narrativa y la identidad numérica (DeGrazia, 2005 y Rodríguez López, 2016). La identidad narrativa se refiere fundamentalmente a lo que somos como personas y lo que nos distingue de los demás; el modo en que nos hemos ido haciendo a nosotros mismos y, por tanto, el modo en que nos entendemos y la imagen que hemos forjado de nuestra personalidad. Sería la respuesta a la pregunta ‘¿quién soy yo?’. La identidad numérica, en cambio, es lo que nos hace ser la misma persona a través del tiempo, pese a ciertos cambios que hayamos podido experimentar; aquello que hace que no desaparezcamos debido a los cambios. Aquí nos centraremos en este segundo sentido de la identidad personal, que es el que más interesa en relación con el problema del volcado de la mente, aunque incluiremos el mantenimiento de la identidad narrativa como una de las propuestas posibles sobre la identidad numérica.¹

3. Análisis del mantenimiento de la identidad personal desde las diferentes perspectivas

La discusión filosófica sobre los criterios para determinar cuándo y cómo se mantiene la identidad de una persona a lo largo del tiempo tiene ya una larga historia. En lo que aquí nos concierne, bastará con señalar los principales criterios que han sido propuestos acerca de la permanencia de la identidad, aun sabiendo que todos tienen problemas y han sido sometidos a críticas por los que defienden posiciones rivales.² Son los siguientes:

- *Continuidad corporal o física.* La permanencia de la identidad tiene como condición necesaria (aunque no suficiente) la permanencia de la continuidad espacio-temporal del cuerpo, o, al menos, del cerebro. Esta sería la posición de los que defienden, por ejemplo, que la consciencia (y sin ella no podríamos hablar de identidad personal) es un fenómeno típicamente biológico, como John Searle (1980, 2008), Jerry Fodor

(2000) o Massimo Piuglicci (2014). Sus críticos suelen acusarla de biologicismo, es decir, de presuponer sin fundamento empírico suficiente que solo una entidad biológica pueda albergar procesos mentales conscientes.

- *Continuidad o conectividad causal psicológica.* El mantenimiento de la identidad radica en la persistencia de la consciencia o conexión de la memoria y de otros rasgos psicológicos, como la personalidad, los intereses y las capacidades sensoriales y cognitivas. Los defensores más conocidos de esta posición son John Locke (1690/1998) (que se centró solo en la memoria), Derek Parfit (2004b) y John Perry (2002). Tal como describe esta posición Eric Olson, “[e]res ese ser futuro que en cierto sentido hereda de ti sus características mentales (creencias, recuerdos, preferencias, la capacidad de pensamiento racional, etc.) y eres ese ser del pasado cuyas características mentales has heredado de esta manera.” (Olson, 2021). Es de suponer que muchos de los que defienden la idea de que la identidad personal consiste en un patrón de procesos mentales, como Hans Moravec (1988, p. 117) o Ray Kurzweil (2012, pp. 371 y ss.), aceptarían que ese patrón es sobre todo una estructura concreta de recuerdos y rasgos psicológicos como los mencionados (Bamford y Danaher, 2017).
- *Ambas cosas.* Normalmente, los que defienden que la continuidad corporal o, para ser más exactos, la continuidad cerebral es condición necesaria, pero no suficiente, para el mantenimiento de la identidad personal, como John Searle o Massimo Piuglicci, asumen que para que se mantenga dicha identidad es necesario añadir también la preservación de la identidad psicológica en el sentido descrito (Piuglicci, 2014, p. 127). Es decir, que buena parte de los que defienden la primera posición descrita, en realidad defenderían más bien esta otra conjunta.
- *Continuidad vital (o de las propiedades animales).* Según esta tesis, solo la preservación de la vida del sujeto, de sus constantes vitales, de su homeostasis como sistema biológico, puede mantener la identidad personal. No debería haber, pues, ninguna discontinuidad en sus procesos vitales para que se dé ese mantenimiento. Si su cuerpo muere, entonces la identidad personal desaparece.
- *Continuidad narrativa.* En la descripción que hace Olson (2021), “en términos generales, un ser pasado es usted solo si ahora tiene narraciones del tipo correcto que lo identifican con ese ser como era entonces”. La preservación de la identidad depende, pues, de la posibilidad de establecer una narración coherente y con sentido que enlace los distintos momentos temporales del individuo.
- *Continuidad del alma o del espíritu.* Es la posición de los que creen que la mente se identifica con, o se sustenta sobre, un sustrato inmaterial o espiritual que puede mantener su existencia tras la muerte. Por lo general, esta concepción se apoya en convicciones religiosas o filosóficas sobre la inmortalidad del alma.
- *Anticriterialismo.* Según esta posición, que es la de los más escépticos en este debate, simplemente no hay criterios no triviales sobre el mantenimiento en el tiempo de la identidad personal que

funcionen como condiciones necesarias y suficientes para determinar ese mantenimiento (Merricks, 1998, Langford, 2017). No podemos decidir la cuestión mediante este tipo de criterios simples que acabamos de presentar, aunque esto no significa que no podamos hablar del mantenimiento de la identidad personal en circunstancias concretas por criterios cotidianos que no tienen el peso metafísico con el que se suele cargar a los criterios propuestos.

No pretendo ser exhaustivo en este recuento, pero creo que las posiciones mencionadas son las principales en el debate. ¿Cómo afectaría el volcado de la mente al mantenimiento de la identidad personal desde cada una de ellas? Veámoslo. Parece claro que, desde la perspectiva de la continuidad corporal o física como criterio del mantenimiento de la identidad personal, ésta no se mantendría después del volcado de la mente en una máquina, puesto que precisamente el volcado significa una ruptura de esa continuidad física, al sustituirse un cuerpo biológico por un cuerpo artificial, constituido sustancialmente por chips de silicio o los equivalentes funcionales que puedan fabricarse en el futuro. Esto podría cuestionarse en el caso del volcado gradual (sustitución una a una de neuronas naturales por artificiales), puesto que parecería que en todo momento la continuidad física (cerebral) se mantiene. Sin embargo, si se acepta la tesis de que la consciencia es un producto biológico, ese volcado gradual podría constituir una pérdida gradual de la consciencia, es decir, algo parecido a la creación de uno de esos zombis hipotéticos que tanto han ocupado la atención de los filósofos de la mente, puesto que se estaría cambiando gradualmente un cerebro biológico por un “cerebro” no biológico. Y, si esto fuera así, no podrían establecerse procedimientos empíricos para determinarlo, puesto que, por hipótesis, los zombis filosóficos, aunque tengan soporte no biológico, son indistinguibles en su comportamiento de las personas conscientes, y si les preguntamos si son conscientes, dirían que sí. Claro que, para aceptar esto, hemos de asumir que la consciencia no es un mero epifenómeno y, además, dar por zanjadas otras dificultades filosóficas que la hipótesis de los zombis trae consigo, lo que es mucho decir (Falk, 2022).

Desde la perspectiva de la continuidad o conectividad causal psicológica, lo que importa, como hemos dicho, es si se heredan los rasgos mentales básicos, como la memoria, las preferencias, el raciocinio, etc. Pues bien, creo que sería discutible, según las circunstancias y procedimientos, si se mantiene o no la identidad personal después del volcado desde este enfoque. Habría que precisar los detalles en cada caso, que aquí serían enormemente relevantes. No en todos los casos imaginables de volcado se heredarían esos rasgos mentales, y hay ejemplos, como los de la mente escindida, en los que esos rasgos serían heredados por dos sujetos distintos. Y algo similar podría decirse si se acepta la realización de varias copias (el conocido como “problema de la reduplicación” (Kind, 2015, cap. 3)). Supongamos que el volcado es exitoso y hago una copia perfecta de mi mente mientras mi cuerpo es destruido. El éxito de esta tecnología debería permitir en principio hacer también dos o más copias de mi mente si ese hubiera sido mi deseo. Suponiendo que la continuidad psicológica se mantiene en todas, ¿cuál de esas copias soy yo? Atendiendo a este problema, podría decirse en general

que ningún volcado no-destructivo puede preservar la identidad, puesto que el original permanecería junto a la copia.³

Incluso aceptando la idea de la identidad-patrón, es decir, la idea de que mi identidad personal consiste en un patrón determinado de procesos mentales, si se añade a esto una concepción funcionalista de esos procesos mentales, tampoco podríamos decir que se mantiene la identidad personal en casos así si nos tomamos en serio el compromiso materialista del funcionalismo, que identifica un estado mental con un estado físico concreto en un sistema físico particular, ya sea un cerebro o una máquina, (*token-token identity*). Puesto que los estados físicos de mi cerebro no serían los estados físicos de la máquina, aunque el patrón funcional fuera el mismo, no podría decirse que los estados mentales correspondientes sean idénticos (Diéguez, 2017, pp. 102-103). En cuanto al volcado gradual, de nuevo se plantearía la duda de si con él no se estaría destruyendo gradualmente la consciencia del sujeto o cambiando gradualmente su identidad. El que una persona melnuda pierda un pelo no lo convierte en calvo, tampoco el que pierda dos, tres o muchos más, pero podemos estar seguros de que a partir de cierto momento ya no es un melnudo, sino un calvo. En el volcado gradual podríamos estar cambiando también gradualmente, pero de forma inexorable, los procesos mentales del individuo, de modo que los cambios podrían ser al final tan grandes que ya no podría decirse que el individuo habría heredado los rasgos mentales del sujeto inicial.

Desde la perspectiva de la postura que combina las dos anteriores, la continuidad física y la psicológica, no se mantendría la identidad personal, puesto que, al menos, faltaría la continuidad física o corporal. En el caso del volcado gradual se sumarían las dudas ya expresadas en ambas.

Desde la perspectiva de la continuidad vital, el volcado de la mente, al implicar la desaparición del organismo biológico, sustituido por una entidad no biológica, no preservaría la identidad. E incluso, si el organismo original no fuera destruido y su cuerpo se mantuviera con vida, supuestamente la mente ya no estaría en ese cuerpo, sino en la máquina, por lo que esa vida sería independiente del mantenimiento de dicha mente y se habría producido una ruptura vital.

En cambio, desde la perspectiva de la continuidad narrativa, sí cabría la posibilidad en principio del mantenimiento de la identidad personal tras el volcado, puesto que podría suponerse que el sujeto es capaz de incluir todo el proceso en una narrativa personal coherente que conecte a la entidad resultante del volcado con la persona anterior al volcado. La cuestión es si realmente esa suposición es plausible en estas circunstancias, teniendo en cuenta el cambio tan radical que implicaría un volcado, es decir, la cuestión es si podría darse a sí misma esa continuidad narrativa coherente una persona que hubiera experimentado el volcado y se hubiera transformado en una máquina de longevidad indefinida y de capacidades muy alejadas de las que tenía en su existencia previa. No puede darse por seguro que fuera así en todos los casos. Es muy dudoso que mi mente en un cuerpo muy distinto, no digamos ya en un cuerpo de metal, plástico y silicio, siguiera siendo yo, en el sentido de que pudiera conectar de forma narrativa los acontecimientos de mi vida pasada, incluyendo las posibilidades que mi cuerpo biológico me abría, con los acontecimientos de mi nueva "vida" en el nuevo cuerpo artificial. De hecho, ante rupturas narrativas de menor calado, como la que puede darse en una persona que fue un traficante de drogas en su juventud, pero

luego se rehabilitó, formó una familia y se convirtió en un ciudadano respetable, solemos pensar que tal persona ya no es la misma que antes, al menos en el sentido importante de que ya no deberían exigírsele responsabilidades penales no cumplidas por los delitos de juventud (Rodríguez López, 2016). En este contexto al menos habría que prestarle oídos a los defensores de la idea de la mente extendida, que el propio Chalmers ha defendido (Clark y Chalmers, 2011). Mis procesos mentales en una máquina, que constituye una extensión de los mismos con capacidades muy diferentes de las que puede proporcionar un cuerpo biológico, difícilmente podrían ser considerados narrativamente como los de la persona que fui mientras tenía un cuerpo biológico.

Desde la perspectiva de la continuidad del alma o del espíritu, podría asumirse que se diera el mantenimiento de la identidad personal tras el volcado, a no ser que partamos de la tesis de que el alma no existe o de que una máquina no puede tener alma mientras que un cuerpo biológico sí puede. El defensor de esta propuesta a buen seguro argumentará que nada impide que, si un cuerpo puede tener un alma, también una máquina podría tenerla, e incluso que pueda ser esa misma alma trasladada de soporte físico. Obviamente, aceptar este argumento implicaría aceptar la existencia del alma, y, además, aceptarla enmarcada en un dualismo radical que la considere como algo separable del cuerpo. Creo que puede decirse con cierta seguridad que, a estas alturas del debate filosófico, un dualismo semejante no puede reclamar una amplia aceptación. Pero, incluso si se aceptara tal dualismo, un problema adicional sería que, en el volcado, la copia de la información cerebral no tendría por qué darnos el alma de la persona (Walker, 2014). Desde luego, ninguna persona que acepte la existencia del alma por motivos religiosos la identificará con información contenida en el cerebro.

Finalmente, desde la perspectiva del anticriterialismo, podría defenderse el mantenimiento de la identidad personal después del volcado, puesto que, ante la falta de criterios necesarios y suficientes para determinar si se mantiene la identidad a lo largo del tiempo, no pueden excluirse “saltos” y fisiones mentales que, pese a todo, la preserven.

Así pues, resumiendo lo dicho, solo las perspectivas de la permanencia de la identidad como continuidad psicológica, como continuidad del alma, como continuidad narrativa o la perspectiva anticriterialista dan pie para aceptar la posibilidad de que tras el volcado de una mente en una máquina el resultado conservara la identidad personal del individuo cuya mente fue volcada. No obstante, como hemos visto, en todos estos casos, excepto en el último, surgen dudas al respecto dependiendo de las condiciones concretas del volcado.

4. Incluso así, el volcado podría ser deseable por otros motivos

Cabe, no obstante, mantener posiciones favorables al volcado incluso aceptando esta conclusión pesimista. Ya hemos visto que esa sería, por ejemplo, la posición de Parfit, para quien cuenta solo la supervivencia (con independencia de la identidad). Podría aducirse, por poner otro ejemplo, que, aunque el volcado de la mente en un ordenador no preservara la identidad personal, podría ser algo deseable para muchas personas si el proceso generara una entidad personal que, al menos en algunos aspectos centrales, mantuviera cierta continuidad con el sujeto original, o preservara algunas cosas importantes del mismo, en el sentido de que

fuera una copia de dicho sujeto capaz de desempeñar funciones similares a éste ante los seres queridos, o capaz de mantener algunas de sus cualidades valiosas. Esta es una posibilidad que discuten con brillantez y detalle Corabi y Schneider (2014), pero no está muy claro, en mi opinión, que a la hora de la verdad el volcado por estos motivos contara con mucha aceptación real, dado que se asume la pérdida de la identidad, con lo cual el sujeto resultante no sería el que tenía esas cualidades que se preservan, aunque, claro está, esto solo podría saberse con seguridad mediante estudios empíricos una vez que algo así fuera posible.

Esta posibilidad plantea una situación análoga a la del mantenimiento de la identidad de tipo (*type*), frente a la identidad de ejemplar (*token*), que, según Walker (2014), podría permitirnos diferenciar a los que defienden la teoría psicológica frente a los que defienden la teoría somática (física) de la identidad. Supongamos que tengo en mi ordenador el texto de la obra de Cervantes *Rinconete y Cortadillo* e imprimo una copia en papel. Al día siguiente un movimiento torpe hace que se derrame el café sobre ella, de modo que la quemo en la chimenea e imprimo una nueva copia. Esa nueva copia no es la misma que la anterior, en el sentido de que son distintos ejemplares (*tokens*), físicamente independientes, de un mismo texto. En cambio, podría decirse que se ha mantenido en el proceso la identidad de la obra (como tipo). La nueva copia, en efecto, sigue siendo el texto de *Rinconete y Cortadillo*. Si la teoría psicológica de la permanencia de la identidad personal tuviera razón, podríamos decir que una copia de mi mente en una máquina no sería idéntica a mí en el sentido de la identidad de ejemplar (no mantendría la continuidad material o corporal), pero sí lo sería en el sentido de la identidad de tipo (sería un ejemplar nuevo de mi mente entendida como tipo). El problema aquí es si esa identidad de tipo mantiene lo suficiente de nuestra identidad personal como para hacer deseable el volcado. No resulta fácil aceptar tal cosa. Esa identidad de tipo parece presuponer que la mente es un patrón de información que puede ser almacenado o transportado de diferentes maneras, un ‘invariante organizacional’ en el sentido de Chalmers, y eso es precisamente lo que está en cuestión en todo este debate.

Por último, habrá quien, incluso aceptando que el volcado de la mente en una máquina suponga la pérdida de la identidad personal, estaría dispuesto a arriesgarse a llevarlo a cabo, porque juzgará que la auténtica identidad que le debe caracterizar ha de ser la futura identidad personal que pueda proporcionarle su existencia en la máquina, su verdadera identidad (así lo cree esa persona al menos) es la que está por llegar cuando su existencia sea en una máquina. Será esa identidad nueva la que le constituirá como el individuo que realmente quiere ser, aunque sea a costa de la desaparición del que ahora es, al que probablemente verá como un ente inauténtico en transición hacia su verdadera plenitud. Es también una posibilidad que no cabe descartar, pero sigue siendo lo más parecido a una muerte, por mucho que algunos afamados transhumanistas la defiendan.

5. Conclusiones

Pese al carácter especulativo que inevitablemente tiene cualquier reflexión sobre la posibilidad del mantenimiento de la identidad personal tras el volcado de la mente en una máquina, dado que ni siquiera podemos afirmar con la más mínima seguridad que tal volcado sea alguna vez tecnológicamente posible, puede

sostenerse que dicho mantenimiento es sumamente problemático analizado desde las principales perspectivas teóricas que pueden adoptarse acerca de la cuestión de en qué consistiría en realidad el mantenimiento de la identidad personal. Incluso en los enfoques más favorables a la idea de que se mantendría la identidad personal tras el volcado, cabe concebir situaciones en las que esto sería muy cuestionable.

Referencias bibliográficas

- Azevedo, F. A. C. *et al.* (2009). Equal Numbers of Neuronal and Nonneuronal Cells Make the Human Brain an Isometrically Scaled-up Primate Brain. *J Comp Neurol*, 513(5), 532-541.
- Bamford, S. y Danaher, J. (2017). Transfer of Personality to a Synthetic Human (Mind Uploading) and the Social Construction of Identity. *Journal of Consciousness Studies*, 24(11-12), 6-30.
- Carroll, J. W. y Markosian, N. (2010). *An Introduction to Metaphysics*. Cambridge: Cambridge University Press.
- Chalmers, D. (2010). The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies*, 17(9-10), 7-65.
- Chalmers, D. (2014). Uploading: A Philosophical Analysis. En R. Blackford y D. Broderick (eds.), *Intelligence Unbound. The Future of Uploaded and Machine Minds* (pp. 102-130). Malden, MA: Wiley Blackwell.
- Clark, A. y Chalmers, D. (2011). *La mente extendida*. Oviedo: KRK Ediciones.
- Corabi, J. y Schneider, S. (2014). If You Upload, Will You Survive? En R. Blackford y D. Broderick (eds.), *Intelligence Unbound. The Future of Uploaded and Machine Minds* (pp. 131-145). Malden, MA: Wiley Blackwell.
- DeGrazia, D. (2005). *Human Identity and Bioethics*. Cambridge: Cambridge University Press.
- Diéguez, A. (2017). *Transhumanismo. La búsqueda tecnológica del mejoramiento humano*. Barcelona: Herder.
- Falk, D. (2022). The Philosopher's Zombie. *Aeon*, 4 de febrero. URL: <https://aeon.co/essays/what-can-the-zombie-argument-say-about-human-consciousness>
- Fodor, J. (2000). *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge, MA: MIT Press.
- Häggström, O. (2021). Aspects of Mind-Uploading. En W. Hofkitchner y H.-J. Kreowski (eds.), *Transhumanism: The Proper Guide to a Posthuman Condition or a Dangerous Idea?* (pp.3-20). Cham: Springer.
- Hopkins, P. D. (2012). Why Uploading Will Not Work, or, the Ghosts Haunting Transhumanism. *International Journal of Machine Consciousness*, 4(1), 229-243.
- Kind, A. (2015). *Persons and Personal Identity*. Cambridge: Polity Press.
- Kurzweil, R. (2012). *La singularidad está cerca. Cuando los humanos trascendamos la biología*. Berlín: Lola Books.
- Langford, S. (2017). A Defence of Anti-Criterialism. *Canadian Journal of Philosophy*, 47(5), 613-630.
- Locke, J. (1690/1998). *An Essay Concerning Human Understanding*. London: Penguin Classics.
- Macdonald, C. (2005). *Varieties of Things Foundations of Contemporary Metaphysics*. Oxford: Blackwell.

- Moravec, H. (1988). *Mind Children: The Future of Robots and Human Intelligence*. Cambridge, MA: Harvard University Press.
- Merricks, T. (1998). There Are No Criteria of Identity Over Time. *Noûs*, 32(1), 106-124.
- Niconiconi (2021). Whole Brain Emulation: No Progress on *C. elegans* after 10 Years. *LessWrong*. <https://www.lesswrong.com/posts/mHqQxwKuzZS69CXX5/whole-brain-emulation-no-progress-on-c-elegans-after-10-years>
- Olson, E. T. (2021). Personal Identity. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.). <https://plato.stanford.edu/archives/spr2021/entries/identity-personal/>
- Ortega y Gasset, J. (1964). Goya. En *Obras Completas*, VII. Madrid: Revista de Occidente.
- Parfit, D. (2004a). *Personas, racionalidad y tiempo*. Madrid: Síntesis.
- Parfit, D. (2004b). *Razones y personas*. Madrid: Antonio Machado Libros.
- Perry, J. (2002). *Identity, Personal Identity, and the Self*. Indianapolis: Hackett Publishing Company.
- Pigliucci, M. (2014). Mind Uploading: A Philosophical Counter-Analysis. En R. Blackford y D. Broderick (eds.), *Intelligence Unbound. The Future of Uploaded and Machine Minds* (pp. 117-130). Malden, MA: Wiley Blackwell.
- Rodríguez López, B. (2016). Cuestión de identidad. Sobre el uso de las tecnologías convergentes para la mejora humana. En C. M. Romero Casabona, (ed.) *Tecnologías convergentes: desafíos éticos y jurídicos* (pp. 7-31). Granada: Comares.
- Searle, J. (1980). Minds, brains and programs. *Behavioral and Brain Sciences*, 3, 417-457.
- Searle, J. (2008). Biological naturalism. En M. Velmans and S. Schneider, (eds.), *The Blackwell Companion to Consciousness* (pp. 325-334). Malden, MA: Blackwell.
- von Bartheld, Ch. S. (2018). Myths and Truths about the Cellular Composition of the Human Brain: A Review of Influential Concepts. *J Chem Neuroanat.*, 93, 2-15.
- Walker, M. (2014). Personal Identity and Uploading. En R. Blackford y D. Broderick (eds.), *Intelligence Unbound. The Future of Uploaded and Machine Minds* (pp. 161-177). Malden, MA: Wiley Blackwell. Publicado originalmente en *Journal of Evolution and Technology*, 22(1), 2011, 37-51. <https://jetpress.org/v22/walker.htm>
- Zamora Bonilla, J. (2019) Por qué es casi seguro que tu mente no está (ni estará nunca) en un ordenador. *Pasajes*, 19-27.

Notas

- 1 Una buena introducción a las diferentes teorías sobre la identidad personal en sentido numérico y sus correspondientes problemas puede verse en Macdonald (2005, cap. 4) y Carroll y Markosian (2010, cap. 5).
- 2 Entre las propuestas que quedan fuera de nuestra consideración por no haber tenido una repercusión amplia, pese a su interés teórico, está la del mantenimiento de la identidad personal como construcción social, según la cual es la relación con los otros, su mirada hacia nosotros, sus creencias sobre nosotros, lo que constituye nuestra identidad personal, al menos en una importante medida. Una defensa de este planteamiento puede verse en Bamford y Danaher (2017). Estos autores concluyen que, bajo esta interpretación de la identidad personal, sería posible su mantenimiento después del volcado, “pero su éxito estaría determinado en última instancia por factores sociales”.

- 3 Y, en tal caso, tiene sentido plantearse, como hacen muchos críticos del volcado de la mente, qué relevancia puede tener para la identidad personal de una mente volcada, una vez que ya lo está, si se destruye o no el original. Como dice Hopkins, “la destrucción de su «viejo» cuerpo es solo una cortina de humo para oscurecer el hecho de que la similitud psicológica exacta entre usted y su copia no es identidad.” (Hopkins 2012, p. 242).