

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN FILOSOFÍA DE LA CIENCIA

FILOSOFÍA DE LAS CIENCIAS COGNITIVAS

UN ESTUDIO DE LA CONSCIENCIA A TRAVÉS DE LA PSICOPATOLOGÍA

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTORA EN FILOSOFÍA DE LA CIENCIA

PRESENTA:

ANDREA MONZERRAT GUERRERO ROMERO

TUTORES PRINCIPALES:

Dr. Santiago Echeverri Saldarriaga (IIF, UNAM)

Dr. Miguel Ángel Sebastián González (IIF, UNAM)

MIEMBROS DEL COMITÉ TUTOR:

Dr. Alejandro Vázquez del Mercado Hernández (FFyL, UNAM)

Dr. Esteban Withrington (IIF, UNAM)

Dr. Pedro Stepanenko Gutiérrez (IIF, UNAM)

Ciudad de México, octubre del 2026

A mi familia,
Martín, Rocío, Alejandro y Luca

y

A mis amigos,
Axel, Berenice, Dalia y Román

Love is to share, mine is for you
S. Tellier

Tabla de contenidos

Agradecimientos……4

Introducción……...5

Capítulo 1. La teoría gradualista……...14

1.1. Introducción……...14
1.2. La subjetividad de la experiencia y el principio subjetivo…….16
1.3. La teoría gradualista……...18
1.4. Dificultades interpretativas……21
1.4.1. Criterios de gradabilidad……...22
1.4.2. La interpretación gradualista de los reportes……22
1.5. Dificultades teóricas……...30
1.5.1. Diferentes concepciones de la subjetividad y características constitutivas…….31
1.5.2. Criterios de gradabilidad para la subjetividad de la experiencia……..36
1.6. Dificultades con la evidencia empírica……..42
1.7. Conclusión……..44

Capítulo 2. La teoría negacionista……47

2.1. Introducción……...47
2.2. La teoría negacionista……49
2.3. Dificultades interpretativas: la interpretación negacionista de los reportes……...52
2.4. Dificultades explicativas respecto al acceso……..62
2.4.1. Acceso sin fenomenología: el caso de la súper visión ciega……62
2.4.2. Dificultades a partir del caso hipotético de Block……65

2.5. Dificultades con la evidencia empírica……..68
2.5.1. Consciencia de acceso y evidencia empírica…….69
2.5.2. Consciencia fenoménica y evidencia empírica……75
2.6. Conclusión……..78

Capítulo 3. La teoría afectivista…….81

3.1. Introducción……...81
3.2. Teorías afectivistas: antecedentes……..82
3.3. La teoría afectivista y sus dimensiones afectivas:
valencia y activación/excitación……...84
3.4. La interpretación afectivista de los reportes……..91
3.5. Evidencia empírica……...119
3.6. Conclusión……122

Conclusiones……..126

Apéndice de reportes…….133

Referencias…….136

Agradecimientos

A la secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por la beca otorgada y, sin la cual, no habría podido realizar mis estudios de doctorado.

Esta investigación contó con el apoyo de los proyectos PAPIIT "Escepticismo y racionalidad" (IA400124) y "Dimensiones de la justificación epistémica" (IN403726). Agradezco a la DGAPA-UNAM por la beca recibida y al responsable de estos proyectos, el Dr. Santiago Echeverri.

A los miembros de mi comité tutor: Esteban Withrington, Alejandro Vázquez y Pedro Stepanenko. Sus comentarios y sugerencias contribuyeron a que esta tesis mejorara a lo largo de los años.

A mis tutores: Miguel Ángel Sebastián y Santiago Echeverri. Agradezco su apoyo incondicional y el ánimo que me proporcionaron en todos estos años que trabajé con ustedes. Asimismo, agradezco los espacios creados por ustedes para poder desarrollar y discutir mis ideas con mis compañeros de estudio, como el laboratorio de filosofía de las ciencias cognitivas y el grupo de investigación Experiencia, Razón y Acción.

Por último, a mis familiares y amigos. Su apoyo y cariño me mantuvieron firme cuando pasé por momentos difíciles.

Introducción

De acuerdo al *Manual Diagnóstico y Estadístico de los Trastornos Mentales* (DSM por sus siglas en inglés), la despersonalización y la desrealización son trastornos disociativos que consisten, respectivamente, en experiencias o sensaciones de "irrealidad o extrañeza, o un distanciamiento de uno mismo en general o de aspectos del yo" [DSM-V, 2014, p. 302] y en experiencias o sensaciones de irrealidad o distanciamiento respecto al mundo y/o los objetos del mundo.

Estas experiencias son consideradas como *anómalas* y se reflejan en los reportes de los pacientes que las padecen. Por ejemplo, pueden reportar que no existen (o que sienten como si estuvieran muertos), que sus estados (pensamientos, emociones, sensaciones) no parecen genuinos o que se sienten desapegados de ellos, que sus partes corporales no se sienten como si fueran suyas o también pueden expresar que "el entorno puede verse como artificial, sin color o sin vida" [Ibíd., p. 303].

Hay un debate respecto a si la despersonalización y la desrealización son trastornos o síndromes separados que pueden ocurrir de manera aislada o si la desrealización puede considerarse como un síntoma de la despersonalización. Sierra y Berrios (1998) y Sierra y David (2011) adoptan la segunda concepción, pues afirman que no hay evidencia concluyente de que sean independientes. Además, señalan que la desrealización aislada de la despersonalización "ha sido cuestionada o reportada como extremadamente rara" [Sierra y David, 2011]. Debido a que no hay evidencia concluyente que apunte a que pueden ocurrir de manera aislada, podemos pensar que, si bien pueden ser diferentes, típicamente ocurren simultáneamente. Es decir, los pacientes que padecen de despersonalización típicamente también presentan episodios de desrealización.

Ahora bien, si en estos casos los pacientes están expresando experiencias anómalas, surgen preguntas filosóficas de gran importancia: ¿En qué consiste la anomalía? ¿Qué aspecto o aspectos de las experiencias pueden fallar para que sean experiencias extrañas o anormales?

Un primer grupo de teorías ubican la anomalía en aspectos generales o estructurales de las experiencias. Tales teorías parten del supuesto según el cual las experiencias son

estados fenoménicamente conscientes. De acuerdo con un análisis influyente, un estado es fenoménicamente consciente solo si hay algo que se siente *para el sujeto* estar en dicho estado (Levine, 2001 y Kriegel, 2009). A partir de ahí, es posible derivar lo que se denomina el "principio subjetivo": todo estado fenoménicamente consciente es necesariamente *subjetivo*. Luego, no puede haber experiencias sin subjetividad (Billon y Kriegel, 2015).

Sin embargo, si los pacientes despersonalizados tienen experiencias de irrealidad o distanciamiento de *ellos mismos* o de aspectos del *yo*, así como experiencias que no se sienten *suyas*, podría pensarse que dichas experiencias amenazan el principio subjetivo. Es decir, si se toman al pie de la letra, los reportes de los pacientes podrían mostrar que hay experiencias sin subjetividad. Luego, podría pensarse que la anomalía de las experiencias que reportan los pacientes deriva precisamente de la ausencia de subjetividad. No obstante, quien quiera interpretar la anomalía como derivada de un aspecto estructural o general de la experiencia deberá encontrar una manera de preservar el principio subjetivo. Existen dos maneras de lograr dicho cometido.

Una primera estrategia la ofrece lo que denomino *la teoría gradualista* según la cual las experiencias que reportan los pacientes despersonalizados son anómalas debido a que estas exhiben una subjetividad atenuada. Es decir, la anomalía deriva del aspecto subjetivo de las experiencias pero, en lugar de ausente, este aspecto sólo está disminuido. De esta manera, la teoría no solo interpreta la anomalía de las experiencias, sino que también resguarda el principio subjetivo, pues, si la subjetividad solamente está disminuida, entonces las experiencias de estos pacientes siguen siendo subjetivas, aunque en un grado menor.

Una segunda estrategia ofrece lo que denomino *la teoría negacionista* según la cual algunos estados de los pacientes despersonalizados no son subjetivos, pero, siguiendo el principio subjetivo, si no lo son, tampoco son fenoménicamente conscientes. De esta manera, también resguarda el principio subjetivo. Sin embargo, la teoría plantea que, dado que los pacientes son capaces de reportar estos estados (a los que llama "estados alienados"), tienen acceso a la información que transmiten.

Una cuestión interesante de esta teoría es que, si las experiencias son estados fenoménicamente conscientes y, si presuntamente los estados alienados de estos sujetos no tienen fenomenología propia, entonces los pacientes no estarían reportando experiencias

anómalas. Sin embargo, lo que queremos saber es de dónde viene la anomalía de las experiencias. Para la teoría negacionista esto no representa ningún problema, pues parece suponer que la anomalía deriva precisamente de la ausencia de fenomenología de los estados de los pacientes. Luego, no están reportando propiamente experiencias, sino estados carentes de fenomenología propia. Más aún, parece ser atípico que un sujeto pueda reportar o tener acceso a la información de estados que no son fenoménicamente conscientes. Siendo así, la anomalía de los estados de estos pacientes deriva de estos dos factores.

En este trabajo voy a argumentar que estas dos teorías enfrentan dificultades y defenderé una teoría que puede evitarlas o resolverlas.

Mi argumento es un argumento a la mejor explicación que divido en dos pasos. En primer lugar, muestro que cada una de las teorías estructurales enfrenta dificultades de tres tipos para dar cuenta de la anomalía de las experiencias de los pacientes despersonalizados. Si bien ninguna de esas dificultades refuta por sí sola cada teoría, intentaré mostrar que, tomadas en conjunto, tales dificultades nos dan razones suficientes para buscar una interpretación alternativa de la anomalía de las experiencias que reportan estos pacientes[1]. En segundo lugar, esbozo una teoría alternativa que denomino *la teoría afectivista* que logra interpretar de manera simple y elegante los reportes de los pacientes. A su vez, esta interpretación afectivista nos permite evitar o resolver las dificultades identificadas en el primer paso del argumento.

En términos generales, la teoría afirma que las experiencias anómalas de los pacientes derivan de la ausencia o disminución de emociones o dimensiones afectivas específicas que, generalmente, acompañan a muchos de nuestros estados mentales.

[1] Respecto a la elección de los reportes, estoy considerando muchos que se recolectaron a principios del siglo XX, pues era común en la práctica psiquiátrica europea recopilar reportes de manera fiel y detallada, una práctica que se ha vuelto mucho menos común en la actualidad. Por otro lado, algunos de los reportes más actuales apuntan en la misma dirección que los reportes más antiguos. Por ejemplo, si el lector consulta el Apéndice, podrá apreciar que Shorvon (1946) y Dugas y Moutier (1911), citan reportes que refieren a emociones y, en la actualidad, Simeon y Abugel (2006) y Sierra (2009) también citan reportes referentes a emociones. O, respecto al dolor, Janet (1928) y Sierra y Berrios (2000), citan reportes similares. El lector también puede consultar similitudes respecto a los pensamientos y la sensación de automatización en el Apéndice. Por último, son estos mismos reportes los que las teorías estructurales utilizan para respaldar sus teorías.

Así, en el primer capítulo, voy a mostrar que la teoría gradualista enfrenta *dificultades interpretativas*, *dificultades teóricas* y *dificultades con la evidencia empírica*. Para tal fin, en la sección 1.2, voy a exponer a detalle qué significa que una experiencia sea un estado fenoménicamente consciente y en qué sentido se puede afirmar que este tipo de estado es subjetivo. A partir de esto, podremos ver cómo se motiva el principio subjetivo.

En la sección 1.3, voy a exponer la teoría gradualista y la predicción que se desprende de ella. De manera general, la teoría sostiene que debemos tomar los reportes de los pacientes literalmente y, a partir de dicha sugerencia, predice lo siguiente: si los estados de los pacientes exhiben subjetividad atenuada, entonces estos reportarán que es *como si* sus experiencias no les pertenecieran. Así, al sostener que la subjetividad está atenuada, en lugar de ausente, la teoría resguarda el principio subjetivo.

Sin embargo, en la sección 1.4, voy a señalar que el defensor de la teoría parece entender la expresión "como si" como un marcador que sugiere implícitamente grados de subjetividad. Pero, lo anterior puede ser problemático debido a que dicha expresión suele entenderse de diferentes maneras que no sugieren grados. Así, en la sub-sección 1.4.1, voy a examinar otros criterios lingüísticos que indiquen la presencia de grados. A partir de dichos criterios y de la sugerencia de tomar los reportes de manera literal, en la sub-sección 1.4.2, voy a mostrar que la teoría enfrenta una serie de dificultades interpretativas. En general, la mayoría de los pacientes no reportan que X es como si no les perteneciera. Siendo así, la mayoría de los reportes no confirman la predicción de la teoría. Aunado a este problema, hay reportes en los que no encontramos indicios de grados, lo que sugiere que no hay atenuación de ningún tipo. Otro problema es que, cuando encontramos grados, estos remiten a experiencias afectivas y no es claro cómo se relaciona la atenuación o disminución de la subjetividad con la disminución o atenuación de las experiencias afectivas. Por último, aunque sí hay algunos reportes en los que los pacientes expresan que X es como si no fuera suyo (o como si no le perteneciera), no parecen referirse a ningún estado mental, contrario a lo que afirma la teoría.

A partir de lo anterior, voy a argumentar que, para solucionar estas dificultades y, en consecuencia, para que los reportes puedan confirmar la predicción de la teoría, esta debería

ser enriquecida con hipótesis auxiliares. Pero, las hipótesis más naturales que podrían añadirse, son problemáticas por diferentes razones.

En la sección 1.5, voy a señalar que, aún si la teoría encontrara una forma de solucionar estos problemas, también enfrenta problemas teóricos. En la sub-sección 1.5.1, voy a mostrar que, aunque en diferentes teorías de la consciencia la subjetividad de la experiencia se caracteriza de diferentes maneras, al menos hay un acuerdo en que la subjetividad debe cumplir con dos características constitutivas íntimamente relacionadas para calificar como tal: *pre-reflexividad* y *elusividad*. En la sub-sección 1.5.2, voy a argumentar que no tenemos razones para pensar que la subjetividad de la experiencia es una propiedad gradable debido a que sus características constitutivas no satisfacen los criterios lingüísticos que indican grados y/o no podemos aplicar dichos criterios sin que desparezcan estas características constitutivas.

En la sección 1.6, mostraré que la teoría gradualista también tiene problemas con la evidencia empírica disponible, pues dicha evidencia no respalda lo que postula la teoría y/o es difícil de reconciliar con lo que afirma dicha teoría.

En la sección 1.7, voy a concluir señalando que, si bien estas dificultades no refutan la teoría, tomadas en conjunto son lo suficientemente serias como para buscar una teoría alternativa que no enfrente dichas dificultades.

En el segundo capítulo, voy a presentar una teoría alternativa que, al igual que la gradualista, busca dar cuenta de la anomalía de los estados de los pacientes despersonalizados sin amenazar el principio subjetivo. Esta es la teoría negacionista. No obstante, voy a mostrar que esta teoría enfrenta dificultades similares a la teoría gradualista. En primer lugar, enfrenta *dificultades interpretativas*. En segundo lugar, enfrenta *dificultades explicativas* y, por último, también enfrenta *dificultades con la evidencia empírica disponible*.

En la sección 2.2, voy a presentar en qué consiste la teoría negacionista y las predicciones que arroja. De manera general, las predicciones son las siguientes: si los estados de los pacientes no son subjetivos, entonces no son fenoménicamente conscientes. Sin embargo, si tienen acceso a dichos estados, los sujetos expresarán que no se sienten como si fueran suyos o, alternativamente, si los estados no son subjetivos y, por tanto, son

fenoménicamente inconscientes, pero estos son accesibles, entonces los sujetos expresarán que son conscientes en un sentido (en el sentido de acceso), pero no en otro (en el sentido fenoménico). Cabe señalar que, al igual que la teoría gradualista, la teoría negacionista también sugiere que debemos tomar los reportes de manera literal. No obstante, a diferencia de la teoría gradualista, la negacionista no toma la expresión "como si" como un marcador que indica implícitamente grados de subjetividad. En cambio, la utiliza para señalar que el sujeto no es delirante.

En la sección 2.3, voy a mostrar que la teoría negacionista enfrenta problemas interpretativos. El problema general es que, así como son pocos los pacientes que reportan que X es como si no fuera suyo (además de que, en estos casos, no parecen referirse a estados mentales), son pocos los pacientes que expresan que son conscientes en un sentido y no en otro (o que son conscientes e inconscientes). Por otro lado y relacionado con este último punto, aunque hay pacientes que sí reportan que son conscientes e inconscientes, no es claro que se refieran inequívocamente a los sentidos de consciencia que plantea la teoría. Otro problema es que, muchos de los reportes de los pacientes no sugieren que sus estados no son subjetivos y, por tanto, no tenemos buenas razones para pensar que son fenoménicamente inconscientes. Por último, la teoría no especifica de qué manera se puede acceder a la información de un estado sin recurrir a experiencias o estados fenoménicamente conscientes.

En vista de estos problemas, voy a argumentar que, de manera similar a la teoría gradualista, la negacionista también tendría que enriquecerse con hipótesis auxiliares para que los reportes de los pacientes confirmen las predicciones que arroja la teoría. Sin embargo, las hipótesis más naturales que podrían añadirse, también resultan ser problemáticas.

En la sección 2.4 y, a partir del último problema que se encuentra a nivel interpretativo, voy a señalar que, mientras que la teoría negacionista asume que hay acceso a los estados supuestamente inconscientes (en el sentido fenoménico), no explica cómo se puede acceder a dichos estados. Para hacer visible el problema, en la sub-sección 2.4.1, voy a exponer el caso hipotético al que recurre Block (1995/2007) para mostrar la posibilidad conceptual de acceso sin fenomenología. Este caso es el de *la súper visión ciega*. En la sección 2.4.2, voy a argumentar que, si los negacionistas están tomando las distinciones de acceso y fenomenología de Block, entonces deberíamos esperar que los casos de

despersonalización se expliquen de manera similar al caso hipotético de la súper visión ciega. En particular, la teoría tendría que postular un mecanismo de acceso análogo al que postula Block para explicar cómo se accede al contenido de un estado fenoménicamente inconsciente. Sin embargo, la teoría no ofrece ninguno y, por tanto, no ofrece una concepción del acceso que no explote (o que no recurra a) la fenomenología del estado al que se accede.

En la sección 2.5, voy a mostrar que la teoría también enfrenta problemas con la evidencia empírica disponible. Por un lado, la teoría sugiere que no hay problemas con el acceso y, como he señalado, que los estados no son fenoménicamente conscientes. No obstante, en la sub-sección 2.5.1, a partir de pruebas neurológicas, voy a mostrar que, de hecho, hay problemas con el acceso. Por otro lado, en la sub-sección 2.5.2, voy a mostrar que si bien es cierto que actualmente no tenemos una idea clara respecto a las áreas cerebrales que subyacen a la consciencia fenoménica, podríamos suponer que ciertas áreas involucradas en la generación de ciertos estados individuales (por ejemplo, percepciones o emociones) contribuyen a la fenomenología de dichos estados. Sin embargo, aun si asumimos esto, voy a mostrar que la evidencia no parece converger con lo que plantea la teoría. En otras palabras, la evidencia no parece respaldar la tesis de ausencia de fenomenología.

Por último, en la sección 2.6, voy a concluir señalando que, al igual que con la teoría gradualista, mi objetivo no es refutar la teoría negacionista, sino mostrar que las dificultades mencionadas previamente son lo suficientemente serias para justificar la búsqueda de una teoría alternativa.

En el tercer capítulo, voy a esbozar una teoría afectivista que nos permite evitar o resolver las dificultades identificadas en el primer paso del argumento (capítulos 1 y 2). De manera general, esta teoría sostiene que las anomalías de las experiencias de los pacientes despersonalizados derivan de *déficits afectivos* que pueden manifestarse de diferentes maneras.

Dado que esta idea no es nueva, en la sección 3.2, voy a exponer qué plantean otras teorías afectivistas para dar cuenta de los reportes de estos pacientes. A partir de esto, voy a señalar cómo difieren estas teorías de la teoría que propongo. En términos generales, las teorías afectivistas afirman que los pacientes exhiben déficits afectivos en tanto no hay

respuesta afectiva o emocional hacia personas, objetos y situaciones del mundo que normalmente la causaban.

La teoría afectivista que defiendo, es similar a estas teorías en tanto concede que, en algunos pacientes despersonalizados, hay una ausencia de respuesta afectiva o emocional y, como consecuencia, suelen reportar problemas con su capacidad para experimentar emociones. No obstante, difiere en tanto no se restringe a falta de respuesta afectiva, sino que especifica dos dimensiones afectivas que no solo están presentes en estados afectivos (como emociones o estados de ánimo), sino que suelen acompañar a muchos de nuestros estados mentales (no emocionales) y, al estar ausentes o disminuidas de estos estados, pueden provocar que estos sean extraños o anormales. Estas dimensiones afectivas son la *valencia* y la *activación/excitación fisiológica*.

En la sección 3.3, voy a desarrollar en qué consisten estas dimensiones afectivas específicas, cómo se entienden en relación con las emociones y cómo suelen estar incorporadas en otros estados mentales. A partir de esto, voy a esbozar a detalle la teoría afectivista que propongo y las predicciones que se desprenden de ella. Las predicciones son las siguientes: si hay déficits afectivos, los pacientes reportarán ausencia o disminución de emociones o, alternativamente, si diferentes estados incorporan dimensiones afectivas como la valencia y la activación/excitación fisiológica, al haber disminución o ausencia de dichas dimensiones, los pacientes lo notarán y tratarán de explicar esta ausencia o disminución utilizando expresiones no literales. En este caso, el déficit afectivo consiste, precisamente, en la ausencia o disminución de estas dimensiones.

Ahora bien, esta teoría difiere de las anteriores en tanto no sigue la sugerencia de tomar los reportes de los pacientes de manera literal. Sin embargo, no quiere decir que no se tome en serio lo que dicen los pacientes. En esta misma sección, señalaré que, en el habla cotidiana, es común utilizar analogías o metáforas para explicar cómo nos sentimos (pues muchas veces es difícil explicarlo recurriendo al lenguaje literal) y que, adicionalmente, se ha planteado que los pacientes despersonalizados solo suelen ser capaces de explicar los cambios en sus experiencias recurriendo a estos recursos lingüísticos.

En la sección 3.4, voy a interpretar los reportes que consideran las teorías estructurales a partir de la teoría afectivista. Si bien es cierto que no debemos esperar que literalmente

reporten que la valencia o la activación fisiológica están ausentes o disminuidas, voy a mostrar que sus reportes apuntan a estas dimensiones teóricas o que es plausible que apunten a ellas.

Por otro lado, al interpretar los reportes de los pacientes en términos de déficits afectivos, voy a mostrar cómo evita o resuelve los problemas interpretativos a los que se enfrentan las teorías estructurales. Además, esta interpretación afectivista nos permitirá dar cuenta de otros aspectos de los reportes que ambas teorías pasan por alto. Estas son claras ventajas respecto a ambas teorías.

En la sección 3.5, voy a mostrar que la teoría afectivista evita los problemas con la evidencia empírica a los que se enfrentan las teorías estructurales, pues, de manera general, dicha evidencia sí converge con los postulados de esta teoría. Lo cual también representa una ventaja respecto a las teorías estructurales.

Por último, en la sección 3.6, voy a concluir señalando que, las ventajas a nivel interpretativo, como a nivel empírico, muestran que la teoría afectivista ofrece una mejor interpretación de las anomalías de las experiencias de los pacientes despersonalizados.

Capítulo 1

La teoría gradualista

1.1. Introducción

Como vimos previamente, la despersonalización se caracteriza, en términos generales, como un padecimiento en el que el sujeto puede sentirse irreal o distanciado de sí mismo, así como de aspectos del yo o referentes a él mismo. Por ejemplo, puede sentir que no tiene pensamientos o emociones o que se siente separado de estos estados. De la misma manera, puede sentir *como si* algunas de sus experiencias no fueran suyas[2].

Ante esto, vimos también que podría pensarse que las experiencias anómalas de estos sujetos son causadas por la ausencia de la subjetividad, pues, precisamente parecen apuntar al sujeto mismo o a aspectos del sujeto. Sin embargo, de ser el caso, los reportes de estos pacientes podrían amenazar el principio subjetivo. Es decir, mostrarían que no es verdad que toda experiencia es necesariamente subjetiva.

Para dar cuenta de las experiencias y, por tanto, de los reportes de estos pacientes sin poner en riesgo el principio subjetivo, la teoría gradualista sostiene que las experiencias anómalas de estos pacientes derivan de una subjetividad atenuada[3] (Billon, 2016).

En este capítulo, voy a mostrar que la teoría gradualista enfrenta tres tipos de dificultades para dar cuenta de la anomalía de las experiencias de dichos pacientes. En primer lugar, enfrenta dificultades interpretativas. En segundo lugar, enfrenta dificultades teóricas y, por último, enfrenta dificultades con la evidencia empírica disponible.

Antes de exponer las dificultades mencionadas, en la sección 1.2, voy a desarrollar a detalle qué significa que una experiencia sea un estado fenoménicamente consciente y qué significa que una experiencia sea subjetiva. A partir de esto, podremos ver cómo se motiva el principio subjetivo.

---

[2] Por lo pronto no voy a citar los reportes específicos de estos pacientes. Esto lo haré hasta la sub-sección 1.4.2 cuando plantee dificultades interpretativas para la teoría gradualista. Respecto a la expresión "como si", veremos que algunos pacientes suelen utilizarla para describir cómo se sienten.

[3] Cabe señalar que, de acuerdo al gradualista, en las formas severas de despersonalización, la atenuación de la subjetividad podría conducir a su ausencia (Billon, 2016, p. 33). Sin embargo, esto obligaría a buscar una manera alternativa de preservar el principio subjetivo. Discuto esta estrategia alternativa para preservar dicho principio en el capítulo 2.

En la sección 1.3, voy a presentar la teoría gradualista y la predicción que se desprende de ella. De manera general, la predicción es que, si la subjetividad está atenuada de los estados de los pacientes, entonces estos reportarán que sienten *como si* sus estados no fueran suyos.

En la sección 1.4, voy a señalar que una primera dificultad con la predicción de la teoría, es que el defensor parece interpretar la expresión "como si" como un marcador implícito que indica que la subjetividad está atenuada. No obstante, la expresión suele entenderse de diferentes maneras, lo que dificulta evaluar si, efectivamente, dicha expresión puede indicar grados de subjetividad.

Ante esta dificultad, en la sub-sección 1.4.1, voy a examinar otros criterios lingüísticos que indiquen la presencia de grados. A partir de estos criterios de gradabilidad y de la predicción de la teoría, en la sub-sección 1.4.2, voy a mostrar que la teoría gradualista enfrenta problemas interpretativos pues, en general, la mayoría de los reportes no cumplen la predicción que arroja la teoría. Una segunda dificultad es que, cuando encontramos indicios de grados y/o expresiones gradables en los reportes de los pacientes, estos remiten a estados afectivos y no resulta claro cómo la atenuación de la subjetividad impacta en la atenuación o disminución de un estado afectivo. La tercera dificultad, es que existe una gama de reportes en los que no encontramos indicios de grados, lo que sugiere que no hay atenuación de ningún tipo. Por último, hay reportes en los que sí encontramos explícitamente que los sujetos expresan que X es como si no fuera suyo. Sin embargo, no parecen referirse a ningún estado mental, contrariamente a lo que dice la teoría. En vista de estas dificultades, voy a argumentar que, para que los reportes cumplan la predicción de la teoría, ésta debería enriquecerse con hipótesis auxiliares. Sin embargo, voy a mostrar que las hipótesis más naturales que se podrían añadir, resultan problemáticas.

En la sección 1.5, voy a mostrar que la teoría gradualista también enfrenta dificultades teóricas. Para tal fin, en la sub-sección 1.5.1, voy a mostrar que, pese a haber diferentes concepciones de la subjetividad de la experiencia, es ampliamente aceptado entre los teóricos de la consciencia que la subjetividad debe cumplir con dos características constitutivas para calificar como tal: *pre-reflexividad* y *elusividad*. En la sub-sección 1.5.2, valiéndome de los criterios lingüísticos de la sub-sección 1.4.1, voy a argumentar que dichas características no

satisfacen los criterios de gradabilidad y/o que no podemos aplicarlos sin que desaparezcan sus características constitutivas. Lo que nos da razones para pensar que la subjetividad no admite grados.

En la sección 1.6, voy a mostrar que la teoría gradualista es difícil de reconciliar con la evidencia empírica disponible y/o que la evidencia no respalda lo que postula la teoría.

Por último, en la sección 1.7, ofreceré las conclusiones del capítulo. Si bien estas dificultades pueden no refutar la teoría gradualista, son lo suficientemente serias como para buscar una teoría alternativa que no enfrente dichas dificultades.

1.2. La subjetividad de la experiencia y el principio subjetivo

En las teorías de la consciencia es ampliamente aceptado que una experiencia es un estado fenoménicamente consciente. Esto es, un estado tal que *hay algo que se siente para un sujeto estar en ese estado*. Siendo así, cuando tenemos una experiencia, estamos en un estado fenoménicamente consciente. Por ejemplo, cuando ves un limón verde sobre la mesa, tienes una experiencia de dicho limón, hay algo que se siente para ti ver el limón. Si lo pruebas, tienes otro tipo de experiencia, hay algo que se siente para ti probar su particular sabor.

Es claro que ambas experiencias son distintas; mientras que la primera es visual, la segunda es gustativa. Más aún, autores como Levine (2001) y Kriegel (2009) han argumentado que ambas experiencias difieren por su *carácter cualitativo*[4]. Por ejemplo, por los objetos y/o cualidades de las que nos percatamos al tener la experiencia. En el caso de la experiencia visual del limón, te percatas de cualidades que tienen que ver con la forma y el color. Mientras que si lo pruebas, te percatas de cualidades como el sabor (la acidez, por ejemplo) o la textura (puedes sentir los pequeños gajos de la fruta en la lengua).

No obstante, mientras que las experiencias pueden diferir por su carácter cualitativo, parece que hay algo que todas ellas comparten. Hay algo que se siente *para-ti* ver el color verde del limón, hay algo que se siente *para-ti* probar la acidez o la consistencia de los gajos.

[4] Voy a tomar la noción de "carácter cualitativo" en un sentido general, de manera que no solo entren cualidades de la experiencia, sino objetos e incluso propiedades tales como "pintura mental" (*mental paint*) o "qualia" (Block, 2003). Si bien estas últimas dos propiedades son controvertidas, mi argumento no depende específicamente de qué tipo de propiedades entran bajo el aspecto cualitativo.

Lo mismo sucede con otro tipo de experiencias. Hay algo que se siente *para-ti* tener dolor de muelas o cosquillas en el pie[5].

Quizás una buena manera de conceptualizar esta observación es sostener que, sin importar qué tan diferentes sean las experiencias, estas siempre se presentan como *tuyas* (Billon y Kriegel, 2015; Gallagher, 2000; Zahavi, 2005). De esta manera, es habitual equiparar nociones tales como *para-mi-idad* (la experiencia es para-ti) con nociones como *mi-idad* (la experiencia es tuya). Billon (2016) ejemplifica esto al sostener que la experiencia "tiene algún tipo de mi-idad, se siente mía, y cuando lo pienso, inmediatamente sé que soy yo quien la está experimentando (…) podemos llamar a este aspecto de para-mí *el carácter subjetivo* o *subjetividad de la experiencia*"[6] [2016, p. 26].

Lo que se pretende capturar con estas distinciones es que, al tener una experiencia, no solo nos percatamos de objetos y cualidades, sino de que la experiencia es para uno mismo o de uno mismo y, en esta medida, debemos percatarnos de nosotros mismos de alguna manera. No obstante, como veremos más adelante, la manera en la que el sujeto se percata de él mismo al tener una experiencia suele ser motivo de debate dentro del estudio de la consciencia fenoménica[7]. Por lo pronto, quedémonos con la idea de que, cuando tenemos una experiencia, esta siempre es para uno mismo o de uno mismo.

Este análisis tiene una consecuencia interesante. Si tus experiencias *siempre* son para ti o son tuyas, sin importar qué tan diferentes sean los objetos y/o cualidades de los que te percatas, entonces el carácter subjetivo es constitutivo de todas tus experiencias. Así, no

[5] Algunos autores piensan que no solo los estados perceptivos, emocionales o de placer y dolor son estados fenoménicamente conscientes, sino también estados cognitivos como las creencias y los deseos. Quienes defienden esta postura se comprometen con la existencia de la fenomenología cognitiva. Para algunas discusiones recientes, véase Billon (2016), Chudnoff (2015), Strawson (2011). Aunque la tesis de la fenomenología cognitiva es controvertida, la aceptaré por mor del argumento. Como veremos más adelante, la teoría gradualista también se compromete con la existencia de la fenomenología cognitiva.

[6] En la sub-sección 1.5.1, veremos que hay ciertas objeciones para tratar los términos como sinónimos y, por lo tanto, que suele ser motivo de controversia a qué se refieren los teóricos de la consciencia cuando hablan de carácter subjetivo.

[7] Parece intuitivo que si te percatas de que la experiencia es para ti o tuya, entonces te percatas de ti mismo. Sin embargo, esto suele admitir diferentes lecturas. Relacionado con la nota anterior, esto también puede dificultar a qué se refieren los autores específicamente por carácter subjetivo.

puede haber experiencias sin subjetividad. Billon y Kriegel (2015) resumen esta observación mediante el llamado *principio subjetivo*[8].

Principio subjetivo: Necesariamente, un estado mental M exhibe consciencia fenoménica solo si M exhibe subjetividad [2015, p. 30].

Dado el principio subjetivo, las experiencias o los estados fenoménicamente conscientes siempre exhiben subjetividad en tanto el sujeto se percata de que son para él o de que son suyos. Luego, si como he señalado, los pacientes despersonalizados reportan sensaciones de irrealidad o distanciamiento de ellos mismos (de un sujeto) o de aspectos referentes a ellos, así como sensaciones de que las experiencias no son suyas, entonces, en principio, estos reportes podrían mostrar que el principio subjetivo es falso.

Sin embargo, ya vimos que una manera de resguardar dicho principio es afirmando que, en los casos de despersonalización, la subjetividad no está ausente, sino solamente atenuada. Esta línea de respuesta da lugar a la teoría gradualista que introduciré a continuación.

1.3. La teoría gradualista

La primera cuestión que debemos entender para desarrollar la teoría es qué significa que la subjetividad está atenuada. De acuerdo con Billon, quien defiende esta teoría, cuando esta particular característica o propiedad de la experiencia está atenuada, ya no aparece claramente para el sujeto. Apoyándose en Simeon y Abugel (2006), sostiene que los pacientes pueden expresar que "la sensación del 'yo' no está claramente presente (en lugar de ausente) y que sus estados mentales no les parecen claramente suyos" [Billon, 2016, p. 26].

Ahora bien, es claro que la despersonalización no se restringe a un tipo de experiencia. Es decir, los pacientes que la padecen reportan un rango amplio de experiencias anormales que refieren a diferentes estados, como pensamientos, emociones o sensaciones. Incluso, hemos visto que pueden sentirse irreales.

[8] Algunos autores que han defendido este principio de manera explícita o implícita son los siguientes: Gallagher (2000, 2004), Zahavi (2005, 2011), Zahavi y Kriegel (2015), Kriegel (2004, 2009), Sebastián (2020, 2022), Duncan (2018), Billon (2013), Campbell (1999) y Graham (2010).

En vista de lo anterior, la teoría postula que los diferentes tipos de experiencias anómalas reportadas por estos pacientes difieren dependiendo de *la extensión* que abarca la subjetividad. Por ejemplo, sostiene que si las sensaciones no se sienten claramente propias, el paciente puede reportar que es como si no fueran sus sensaciones, si no siente claramente que son sus pensamientos, entonces expresará que es como si no le pertenecieran y así con cada estado individual. Asimismo, sostiene que la subjetividad no solo puede abarcar diferentes estados individuales o, incluso, múltiples estados, sino que la atenuación de dicha subjetividad puede ser *más intensa o menos intensa*. Así, sostiene lo siguiente:

> Las formas más leves de despersonalización se explicarían por una ligera atenuación del carácter subjetivo (…) o una que se limita a un sub-conjunto relativamente pequeño de los estados de los pacientes (…). Por otro lado, las formas más graves de despersonalización se caracterizarían por una disminución sustancial o incluso una eliminación completa del carácter subjetivo y que no se restringe a un tipo específico de estados mentales, sino que se extiende a una parte significativa de los pensamientos de los que se percata el paciente. Cuando la falta de subjetividad alcanza tal clímax, el paciente dejaría de sentirse propiamente como *él mismo*. No se percataría de él mismo como un 'yo' y empezaría a sentirse irreal." [Billon, 2016, pp. 32-33].

Como podemos observar, la sensación de irrealidad que reportan algunos pacientes, sería un caso en el que múltiples estados del sujeto no se sienten propios y, además, en los que la subjetividad está gravemente disminuida o incluso ausente. No obstante, si hay ausencia de subjetividad, como he señalado en la sección 1.1, esto nos obligaría a buscar una estrategia alternativa para resguardar el principio subjetivo. Esta estrategia la abordaré en el segundo capítulo.

Así, lo que la teoría postula es que la subjetividad se atenúa o disminuye en mayor o menor grado en tanto el sujeto no siente claramente que sus estados son suyos y que la atenuación de la subjetividad puede abarcar un estado específico o múltiples estados.

A partir de todo lo dicho anteriormente, la teoría gradualista arroja una predicción general: Si un estado o múltiples estados exhiben subjetividad atenuada, entonces los pacientes no sentirán que son claramente suyos y, en consecuencia, expresarán que es como si no fueran sus estados o como si no les pertenecieran.

Por último, cabe señalar que, de acuerdo a la teoría gradualista, debemos tomar los reportes de los pacientes *al pie de la letra*. Esta afirmación admite dos lecturas: 1) que debemos tomar los reportes de los pacientes en serio o tomarnos seriamente lo que dicen o 2) que debemos interpretarlos de manera literal. Esto es, que cuando el sujeto dice que X no se siente como si fuera suyo, es porque literalmente no lo siente propio. A partir de la predicción de la teoría, esta parece apostar por la segunda opción[9]. Por supuesto, si se toman de manera literal, también los estaríamos tomando en serio. La apuesta de la teoría es, precisamente, que para tomarlos en serio, debemos tomarlos de manera literal. De otra manera, no podríamos dar sentido a los extraños reportes de estos pacientes o no podríamos comprender las experiencias que comunican los pacientes. Sin embargo, de tomarlos de manera literal, surgiría una seria de dificultades, comenzando por dificultades interpretativas. A continuación presentaré dichas dificultades.

[9] La sugerencia del gradualista de tomar los reportes de manera literal no es explícita. Sin embargo, podemos apreciar dicha sugerencia por las siguientes razones: en primer lugar, uno de los objetivos del gradualista es dar cuenta de los delirios que expresan los pacientes que padecen síndrome de Cotard (una condición que se caracteriza por delirios de negación dirigidos a aspectos del sujeto como a aspectos del mundo externo) a partir de los reportes de los pacientes con despersonalización. Esta estrategia se debe a que se ha puesto en duda que se pueda dar sentido a los delirios que expresan los pacientes con Cotard debido a que es difícil pensar que existan experiencias como aquellas que presuntamente reportan estos pacientes. Ante esto, Billon señala que puede pensarse que los delirios de estos pacientes son el resultado de procesos de razonamiento anormales y/o de una incapacidad de razonar lógicamente. No obstante, Billon recurre a los pacientes despersonalizados para demostrar que las experiencias que reportan los pacientes con Cotard, de hecho, existen, pues los pacientes despersonalizados expresan reportes similares y, según el autor, estos últimos pacientes no son irracionales y/o exhiben una racionalidad normal (véase Billon, 2016, pp. 21-22). Luego, la idea es que los reportes que expresan los pacientes despersonalizados reflejan de manera fiel experiencias que no se sienten suyas. Ahora, si son similares a las de los pacientes con Cotard, entonces estos también padecen experiencias que no se sienten propias (más aún, derivadas de la misma anomalía). La diferencia es que, mientras que los pacientes despersonalizados no las respaldan (aunque no las sienten suyas, aún creen que lo son), los pacientes con Cotard sí lo hacen, lo que da lugar al delirio y/o al aspecto delirante que caracteriza al síndrome de Cotard (véase Billon, 2016, p. 35). Si bien la interpretación de los delirios de los pacientes con Cotard puede ser controvertida, lo que importa para el argumento es que, siguiendo a Billon, los pacientes con despersonalización expresan de manera fiel que X no se siente suyo. De manera similar, en estudios posteriores, Billon (2023) señala que, dado que los pacientes despersonalizados son racionales, "no tenemos razones para sospechar que están mintiendo (…) o que no dicen en serio lo que parecen indicar" [Billon, 2023, p. 318]. Refiriéndose a que no tenemos razones para pensar que no dicen en serio que X no se siente suyo. Por último, en otros estudios, también ha señalado que, si los reportes se toman al pie de la letra (literalmente) "es natural interpretar (que) el paciente no experimenta el estado consciente relevante como suyo" [Billon y Kriegel, 2015, p. 32. Paréntesis mío].

1.4. Dificultades interpretativas

Una primera dificultad que surge a partir de la predicción de la teoría y de la sugerencia de interpretar los reportes literalmente proviene de la expresión "como si" que, presuntamente, utilizarían los pacientes en sus reportes. Esto se debe a que dicha expresión es ambigua. Parece que el defensor de la teoría la toma como un marcador implícito que indica que la subjetividad está atenuada, pues, como hemos visto, la predicción es que, si cierto estado no se siente *claramente suyo* (del paciente), este reportará que es *como si* no le perteneciera[10]. Esto puede reflejarse a partir de la siguiente interpretación: "de hecho, si mis sensaciones corporales no se sintieran *claramente mías*, podría parecerme *como si* no fueran mis sensaciones corporales" [Billon, 2016, p. 27. Énfasis mío].

No obstante, el mismo gradualista señala que, típicamente, la expresión indica que los pacientes despersonalizados, a diferencia de los pacientes que padecen de síndrome de Cotard, no son delirantes y/o que la despersonalización es la versión "como si" (no delirante) del síndrome de Cotard (Billon, 2016, p. 23). Por otro lado, autores como Sierra y David (2011), han señalado que el uso de la expresión puede indicar "una crítica a la pertinencia de la descripción utilizada" [2011, p. 100]. Es decir, el paciente puede cuestionar si la expresión es la adecuada para describir su experiencia.

Ante esta dificultad, debemos encontrar una manera de evaluar si, efectivamente, la expresión "como si" constituye una manera implícita de indicar que la subjetividad de estos pacientes está atenuada o disminuida o, si es la única manera que tenemos disponible para indicar dicha atenuación. Una manera natural de alcanzar este objetivo es al examinar otros marcadores lingüísticos de la presencia de grados y buscar si estos son o no prevalentes en los reportes de los pacientes. Para tal fin, en la sub-sección 1.4.1, introduciré tres criterios lingüísticos ampliamente aceptados en la literatura. A partir de dichos criterios, en la sub-sección 1.4.2, voy a mostrar problemas específicos con la interpretación gradualista de los reportes.

[10] Esto es, puede no reportar explícitamente que no es claramente suyo, pero, al no sentirlo propio, expresa que es como si no lo fuera. Lo que marca implícitamente la atenuación de la subjetividad.

1.4.1. Criterios de gradabilidad

En un estudio influyente, Kennedy (1999, 2007) establece ciertos criterios para determinar si un adjetivo es gradable. Estos criterios son relevantes para el argumento porque hemos aceptado la sugerencia de Billon según la cual deberíamos tomar seriamente (literalmente) los reportes de los pacientes. Dado ese supuesto, si alguien emplea un adjetivo gradable, entonces ese alguien refiere o busca referir a una propiedad gradable. Los criterios son los siguientes:

1. Un adjetivo gradable debe aceptar construcciones comparativas del tipo "X es más alto que Y" o "X tiene más altura que Y". Por ejemplo, "Juan es más alto que Andrés".

2. Los adjetivos gradables deben dar sentido a las preguntas del tipo "¿qué tan alto es X?" o "¿qué tanta altura tiene X?". Por ejemplo, "¿Qué tan alto es Juan?"

3. Los adjetivos gradables deben admitir modificadores de grados como "X es extremadamente alto", "Y es claramente alto" o "Z es ligeramente alto"[11]. Para este criterio podríamos decir que "Juan es extremadamente alto" (para su edad, por ejemplo). Luego, "alto" es un adjetivo gradable y denota una propiedad (la altura) que se presenta en grados.

Ahora, utilizaré estos criterios para mostrar que la teoría gradualista enfrenta serias dificultades para interpretar un gran número de reportes.

1.4.2. La interpretación gradualista de los reportes

La primera dificultad y que estará presente en casi todos los reportes que revisemos, surge del problema de que muy pocos pacientes expresan que X no se siente como si fuera suyo. Siendo así, no solo los reportes de los pacientes no cumplen la predicción de la teoría, sino que, aunado a este problema, podrían poner en duda la sugerencia de que la expresión "como si" es la que indica que la subjetividad está atenuada. La segunda dificultad aparece cuando examinamos reportes en los que se utilizan expresiones gradables: en estos casos, los pacientes se refieren a estados afectivos y no resulta claro cómo la atenuación de la

[11] Estos criterios los he retomado del trabajo de Hawthorne y Logins (2021), quienes los han extraído del trabajo de Kennedy (2007) para analizar si el adjetivo "justificado" puede admitir grados.

subjetividad impacta en la atenuación o disminución de un estado afectivo. La tercera dificultad reside en que hay una gama de reportes donde no encontramos indicios de grados, lo cual sugiere que la anomalía no siempre se refiere a una atenuación de ningún tipo. La cuarta dificultad es que, aunque en algunos reportes sí encontramos la expresión "como si X no me perteneciera", dichos reportes no parecen referir a estados mentales, contrariamente a lo que dice explícitamente la teoría.

Argumentaré que, para que los reportes cumplan la predicción de la teoría gradualista, esta tendría que enriquecer su propuesta con hipótesis o supuestos auxiliares. Evidentemente, no podré demostrar que no existe ninguna hipótesis auxiliar que pueda dar cuenta de los reportes. Lo que sí haré, es mostrar que algunas de las hipótesis más naturales son problemáticas debido a que son poco plausibles, no están respaldadas por razones independientes o son poco específicas.

A continuación, presentaré cada dificultad mencionada previamente, citaré algunos reportes representativos que la ilustran y, al final de cada caso, presentaré algunas de las hipótesis que podrían añadirse para que se cumpla la predicción de la teoría y los problemas que se derivarían de dichas hipótesis[12].

Atenuación, pero en estados afectivos: Como hemos visto, la teoría predice que, si cierto estado del paciente no se siente *claramente suyo*, entonces expresará que es *como si* dicho estado no le perteneciera; además, el gradualista parece interpretar la expresión "como si" como un marcador implícito que indica que el estado del sujeto no se siente claramente suyo. Luego, la expresión sugiere implícitamente grados de subjetividad.

Sin embargo, si consideramos algunos reportes que refieren a estados afectivos, lo primero que podemos notar es que los pacientes no sugieren que sus emociones o sentimientos, por ejemplo, no se sienten como si fueran suyos. En cambio, reportan lo siguiente: "puedo sentirme entumecido de sentimientos, casi vacío por dentro. Odio el hecho

[12] La primera dificultad estará presente tanto en la segunda como en la tercera, así que no considero abordarla por separado. Por esta razón, comenzaré por la atenuación de los estados afectivos. Por otra parte, como señalé en la introducción de este trabajo, los reportes que voy a citar son reportes que la misma teoría utiliza para respaldar su propia teoría. Esto evita elegir reportes que la teoría no ha considerado y que puede perjudicarla injustamente. Aunque considero que otros reportes enfrentarían dificultades similares, no los cito todos por motivos de extensión.

de que no puedo sentir las cosas como solía hacerlo" [Sierra, 2009, p. 27] y "la parte emocional de mi cerebro está muerta… todas mis emociones están aplanadas" [Shorvon, 1946, p. 784]. Como señalé anteriormente, el hecho de que no utilicen la expresión "como si", puede sugerir que dicha expresión no es la adecuada para indicar una atenuación de la subjetividad. Pero, incluso si concediéramos que la indica, no hay nada en los reportes que lo sugiera, pues, precisamente, no están reportando que es como si no fueran sus emociones.

Ahora, si consideremos los criterios lingüísticos de gradabilidad de la sub-sección 1.4.1, podemos notar que, aunque los pacientes utilicen expresiones que son gradables, éstas invariablemente remiten a la atenuación de la afectividad, no de la subjetividad. Como acabamos de ver, están diciendo que se sienten *entumecidos* de sentimientos o que sus emociones están *aplanadas*. Ambas expresiones cumplen con los criterios de gradabilidad.

Por ejemplo, para las construcciones comparativas podemos decir que mi brazo está más entumecido que mi pierna o que un sentimiento está más entumecido que otro y, de la misma manera, que mi ansiedad está más aplanada que el resto de mis emociones. También podemos preguntar "¿qué tan entumecido te sientes o qué tan aplanadas están tus emociones?". Y, por último, podemos contestar a estas preguntas apelando a modificadores de grados como "las siento ligeramente aplanadas" o "me siento claramente o extremadamente entumecido." Por otro lado, parece natural entender la atenuación de los sentimientos o las emociones como una reducción en la intensidad de la emoción sentida.

Si esto es el caso, para que la teoría pueda dar cuenta de estos reportes, tendría que recurrir a una hipótesis adicional según la cual si la subjetividad de una emoción disminuye, entonces la emoción puede verse disminuida en intensidad o, alternativamente, que aunque el sujeto no exprese explícitamente que sus emociones no se sienten claramente suyas, de cualquier manera la subjetividad está atenuada y, por esta razón, sus emociones se sienten disminuidas.

En tratamientos posteriores sobre la subjetividad de la experiencia, el partidario de la teoría de hecho parece suponer lo anterior, pues sostiene que, "una emoción (…) cuyo sentido de pertenencia está parcialmente ausente es, en el mejor de los casos, una emoción atenuada." [Billon, 2023, p. 333]. No obstante, no ofrece ninguna razón independiente de por qué esto sería el caso. Una buena teoría debería ser capaz de ofrecer razones que conecten la relación

entre ambos aspectos. A lo sumo, lo que supone en relación con las emociones y la subjetividad es lo siguiente: "considera un episodio emocional estándar, por ejemplo, un episodio de miedo, e imagina qué sucedería si eliminásemos el carácter subjetivo de la experiencia (…) involucrado en este episodio. Como ya hemos visto, te parecería como si tú mismo no tuvieras miedo" [Billon, 2016, p. 42].

Sin embargo, hay dos problemas con este supuesto. Por una parte, lo que la teoría postula es que la subjetividad está atenuada, pero en la cita anterior el partidario afirma que si la subjetividad *se elimina*, entonces parecería como si no tuvieras miedo. Es decir, en este caso no está considerando una subjetividad atenuada, sino una ausencia de subjetividad y, lo que esta estrategia trata de mostrar (a diferencia de la estrategia alternativa que veremos en el segundo capítulo), es que hay subjetividad, aunque en un menor grado. Por otra parte, ya vimos que los pacientes no reportan nada parecido a lo que se sugiere. Es decir, no reportan que es *como si no* tuvieran miedo o cualquier otra emoción. Ya vimos que reportan que sienten *sus* emociones aplanadas o entumecidas. Luego, siguiendo la sugerencia de la teoría de tomar los reportes literalmente, los pacientes sugieren que tienen miedo (y/o es *su miedo* o cualquier otra emoción) pero éste está disminuido o no es tan intenso como solía serlo antes de su episodio de despersonalización. En resumen, la atenuación de las emociones no parece estar está relacionada con la subjetividad atenuada.

Ausencia de estados, no atenuación: De nuevo, la teoría postula atenuación o grados de subjetividad. No obstante, como ya he señalado, muchos reportes no remiten a la subjetividad de la experiencia y, en general, no reportan estados atenuados o disminuidos. Esto se refleja en el hecho de que no utilizan adjetivos o expresiones gradables en sus reportes.

Por ejemplo, un sujeto expresa lo siguiente: "tengo la sensación de no tener ningún pensamiento" [Sierra y Berrios, 2000, p. 161]. En este caso, la teoría postula que la subjetividad se ha atenuado de los pensamientos del sujeto, pero si esto fuera así, siguiendo la teoría, el sujeto debería expresar que siente c*omo si no fueran sus pensamientos o como si no le pertenecieran.* Además, la expresión "no tengo ningún pensamiento" no indica grados.

En vista de que la teoría no puede dar cuenta del reporte siguiendo sus propios postulados, necesita recurrir a una hipótesis o supuesto adicional. Por ejemplo, podría suponer (de manera similar al caso de las emociones disminuidas) que, aunque el sujeto no

exprese explícitamente que sus pensamientos no son claramente suyos o siente como si no le pertenecieran, de cualquier manera la subjetividad está disminuida y, como consecuencia, se ve disminuida la sensación de tener pensamientos.

Sin embargo, esto resulta implausible; el sujeto no utiliza ninguna expresión o adjetivo gradable en relación con los pensamientos. Por ejemplo, no expresa o ni siquiera sugiere que tiene la sensación de que sus pensamientos están aplanados o que tiene la sensación de tener menos pensamientos o cualquier otra expresión que indique grados. De nuevo, lo que expresa es que tiene la sensación (o quizás la impresión) de no tener *ninguno*.

Por otro lado, algunos sujetos expresan que no tienen emociones o estados de ánimo. Los siguientes reportes ejemplifican este caso: "besar a mi esposo es como besar una mesa, señor. La misma cosa… ni la menor emoción. Nada en la tierra puede emocionarme. Ni mi esposo ni mi hijo…mi corazón no late. No puedo sentir nada." [Dugas y Moutier, 1911, p. 109] y "un estado de nada, sin ningún estado de ánimo, como si estuviera muerto" [Simeon y Abugel, 2006, p. 30].

Respecto al primer reporte, la teoría plantea que, "si la subjetividad se desvanece de las emociones, (los pacientes) podrían sentir como si no tuvieran emociones en absoluto" [Billon, 2016, p. 28. Paréntesis mío]. Sin embargo, esto contradice la predicción original de la teoría. De nuevo, lo que predice es que si se atenúa de un estado (en este caso, de las emociones), entonces el sujeto expresará que es como si no fueran sus emociones, no como si no las tuviera *en absoluto*.

Ahora, la mujer no solo no dice que es como si no fueran sus emociones, sino que tampoco expresa que es *como si* no las tuviera. Precisamente dice que *no tiene ninguna emoción* o que *no puede sentir ninguna emoción*. Veíamos en el primer problema que, con los grados de las emociones, la teoría podía recurrir a la hipótesis según la cual si la subjetividad de una emoción disminuye (y aunque el sujeto no exprese explícitamente que no la siente claramente suya), entonces la emoción podía verse disminuida en intensidad. Sin embargo, no solo no ofrece una razón independiente que pueda respaldar dicha hipótesis; incluso si pudiera ofrecerla o si lograra relacionar la disminución de la subjetividad con la disminución de la emoción, esta hipótesis resultaría inútil en este caso, pues el sujeto no está utilizando expresiones gradables que indiquen que sus emociones están disminuidas.

El mismo problema lo podemos encontrar en el segundo reporte, a saber, “un estado de nada, sin ningún estado de ánimo, como si estuviera muerto”. En este caso, el sujeto tampoco expresa estados de ánimo atenuados, sino ausencia de ellos. Pero, más allá de este problema, la teoría está interpretando dicho reporte como uno que expresa muerte y supone que, cuando el paciente afirma que se siente muerto, se refiere a que se siente mental o fenoménicamente muerto. Esto es, le parece que no tiene estados mentales fenoménicamente conscientes debido a que la subjetividad se ha atenuado de la mayoría o incluso de todos sus estados mentales.

No obstante, hay dos problemas con este supuesto. Por una parte, es implausible suponer que el sujeto quiere decir que se siente fenoménicamente muerto, pues es poco probable que la gente ordinaria (no filósofa) sepa cuál es la diferencia entre tener estados fenoménicamente conscientes e inconscientes. Por otra parte, el reporte claramente apunta a una dimensión afectiva o sugiere que se refiere únicamente a estados afectivos, no a múltiples estados de otro tipo. Además, sugiere que la referencia a la muerte no es literal; es una analogía o metáfora que utiliza para expresar precisamente la deficiencia que tiene para experimentar estados de índole afectivo. Luego, además de que no encontramos indicios de grados, es implausible adoptar este último supuesto.

Por último, hay un reporte en el que la teoría supone que hay una subjetividad atenuada de las emociones. “Parece que estoy caminando en un mundo que reconozco pero que no siento” [Sierra, 2009, p. 32]. Sin embargo, no es claro por qué hay emociones involucradas en esta descripción. La teoría no lo especifica y una buena teoría debería ser capaz de indicar cómo se relacionan las emociones con reconocer el mundo, pero no sentirlo. Además ¿qué podría significar esto? La teoría tampoco lo resuelve.

Independientemente de cómo se podrían resolver estas cuestiones, hay una clara diferencia entre sentir menos o no sentir claramente algo y no sentirlo en absoluto, que es lo que parece sugerir el paciente. Luego, estos reportes ofrecen evidencia en contra de la teoría gradualista.

Falta de referencia directa a estados mentales: La teoría predice que lo que no se siente como si fuera del paciente son estados mentales. Sin embargo, hay pacientes que reportan que no se sienten vivos, que sus partes corporales no se sienten propias o que sienten que sus

movimientos son automáticos. En estos casos, no solo no hay una referencia a la subjetividad, sino que tampoco parecen apuntar a ningún estado mental.

Consideremos el primer tipo de reporte: "no me siento vivo de ninguna manera" [Sierra, 2009, p. 8]. Ante este caso podemos preguntar, ¿cómo es que la subjetividad atenuada puede explicar que no me sienta vivo? Y en relación con esta pregunta ¿de qué tipo de estado se ve disminuida la subjetividad? Según la teoría, si el sujeto dice que no se siente vivo, entonces esto implica que se siente muerto y, al hacer una referencia a la muerte como en el caso pasado del sujeto que afirma que es "como si estuviera muerto", supone también que cuando este paciente afirma que se siente muerto, se refiere a que se siente mental o fenoménicamente muerto debido a que la subjetividad se ha atenuado de todos sus estados mentales (o al menos de la mayoría de ellos). Sin embargo, ya vimos que adoptar este supuesto es poco plausible[13].

Ahora bien, hay algunos reportes donde claramente encontramos la expresión "como si X no me perteneciera", pero, como ya he mencionado, en estos casos los sujetos se refieren a partes corporales. Veamos los reportes: "las partes de mi cuerpo se sienten como si no me pertenecieran" [Sierra y Berrios, 2000, p. 160] y "notaba que mis manos y pies se movían, pero como si no me pertenecieran y se movieran automáticamente" [Sierra, 2009, p. 29].

Respecto al primer reporte, la teoría supone que la subjetividad se ha atenuado de las sensaciones corporales y utiliza dicho supuesto para extraer dos conclusiones: que tales sensaciones no se sienten claramente suyas y que, en consecuencia, el sujeto siente como si sus partes corporales no le pertenecieran. El problema reside en el último paso del razonamiento. Para que la conclusión se siga de lo anterior, parece que deberíamos suponer que las sensaciones corporales fundamentan necesariamente la sensación de pertenencia del cuerpo. No obstante, este supuesto no está independientemente justificado y parece, además, enfrentar una objeción independiente.

Si bien típicamente nuestras sensaciones corporales se sienten (o están localizadas) en nuestro propio cuerpo, hay sujetos que reportan que pueden sentir sensaciones localizadas

[13] De la misma manera, si en el caso anterior el sujeto está utilizando una analogía o metáfora para referirse a la sensación de estar muerto, es posible que en este caso el sujeto también esté recurriendo a estos recursos lingüísticos para expresar algo que difícilmente podría expresar mediante el uso literal del lenguaje.

en el cuerpo de alguien más. La somatoparafrenia es un caso paradigmático. Los pacientes que padecen de esta condición expresan que alguna de sus partes corporales no le pertenece. Además, la mayoría afirma que no sienten nada en la parte corporal que niega que es suya. Sin embargo, algunos reportan que pueden sentir una sensación táctil en una extremidad que no es suya (o que no se siente suya), sino de alguien más (el lector puede consultar los capítulos 1 y 3 en Gennaro, 2015, para una descripción de esta psicopatología).

El punto relevante de estos casos es que, si pueden sentir sensaciones corporales y aun así la parte corporal no se siente suya, entonces podemos pensar que una no fundamenta a la otra o que no hay una relación necesaria entre ellas. Son aspectos o dimensiones diferentes que, si bien pueden ir de la mano, pueden disociarse. Siendo así, aún si concedemos que la subjetividad puede atenuarse de las sensaciones corporales, parece que no tenemos razones suficientes para pensar que esto explique por qué el sujeto siente como si sus partes corporales no le pertenecieran. Quizás el partidario de la teoría gradualista podría encontrar una explicación alternativa de este fenómeno. Sin embargo, esto le obligaría a enriquecer su teoría mucho más allá de lo que ésta afirma de manera explícita.

Por último, en el caso de los movimientos automáticos, la teoría supone que es una intención la que no se siente claramente propia o como si no fuera la intención del sujeto y que esto explica por qué no se siente en control de su cuerpo (Billon, 2016, p. 28). Sin embargo, podemos encontrar dos problemas con este supuesto. En primer lugar, ya vimos que el sujeto también reporta que algunas de sus partes corporales (manos y pies) se sienten como si no le pertenecieran. Si siente que no le pertenecen, ¿por qué no se asume que también hay una atenuación de la subjetividad en las sensaciones corporales? La teoría podría replicar que siente que no le pertenecen porque, aunque nota que se mueven, parece que se mueven automáticamente. Pero, si esto es así, entonces hay una ambigüedad clara respecto a aquello que fundamenta la pertenencia corporal, pues primero sugieren que es una sensación corporal y en este caso una intención.

En segundo lugar, no hay nada en la formulación de la teoría que prediga que es una intención, y no algo diferente, lo que explica la sensación específica de no tener control sobre el cuerpo. Además, es difícil ver qué razones independientes podrían darse para pensar que debería ser una intención y no otro aspecto de la experiencia.

En suma, a partir de la predicción de la teoría y de la sugerencia de tomar los reportes de manera literal, encontramos, en primer lugar, que la expresión crucial "como si" es ambigua. Esto nos llevó a examinar otros marcadores lingüísticos que indiquen grados o disminución de subjetividad. Sin embargo, al examinar varios reportes representativos, no solo encontramos que muchos de los reportes no cumplían la predicción de la teoría y que, aunque algunos la satisfacían, no hacían referencia directa a estados mentales. Pero también encontramos que, al hacer uso de los marcadores lingüísticos, algunos reportes apuntaban a una atenuación de la afectividad, no de la subjetividad y que otro conjunto de reportes no apuntaban a una atenuación de ningún tipo.

En vista de lo anterior, argumenté que para que la teoría pudiera solucionar dichas dificultades, tendría que ser enriquecida con hipótesis auxiliares que resultaban ser poco plausibles, poco específicas o que no se ofrecían razones independientes para adoptarlas. Si bien no he demostrado que estas dificultades son insuperables, creo que sí he mostrado que se trata de dificultades suficientemente serias que solo podrían resolverse al modificar la teoría de manera considerable.

En la siguiente sección examinaré si, pese a estos resultados negativos, podría haber razones teóricas para sostener que la subjetividad admite grados de atenuación. Voy a argumentar que no tenemos razones para pensar que esto es el caso y, por tanto, la teoría también presenta dificultades teóricas.

1.5. Dificultades teóricas

Supongamos que la teoría gradualista es capaz de resolver todos los problemas mencionados en la sección anterior. Sin embargo, podemos preguntarnos si la subjetividad de la experiencia es una característica o propiedad gradable. En esta sección, ofreceré razones para responder negativamente a esta pregunta. En primer lugar, en la sub-sección 1.5.1, voy a mostrar que, pese a haber diferentes concepciones de la subjetividad, suele aceptarse que esta debe cumplir con dos características constitutivas para calificar como tal: pre-reflexividad y elusividad. Se trata de características constitutivas, pues es posible encontrarlas en la mayoría, si no en todos, los análisis existentes de la subjetividad. En segundo lugar, en la sub-sección 1.5.2, argumentaré que no parece posible imaginar casos en los cuales tales características pueden estar atenuadas. Para tal fin, me valdré de los criterios lingüísticos

como instrumentos heurísticos. Mostraré que, cuando logramos aplicar tales criterios a aspectos de la experiencia, tales aspectos siempre se refieren a aspectos no-subjetivos, que denominaré, de manera genérica, aspectos cualitativos. Y, cuando finalmente logramos enfocarnos en la subjetividad, las características previamente mencionadas no cumplen los criterios lingüísticos y/o no podemos aplicarlos sin que desaparezcan sus características constitutivas.

1.5.1. Diferentes concepciones de la subjetividad y características constitutivas

En la sección 1.2, mencioné brevemente que es motivo de debate a qué se refieren los teóricos de la consciencia cuando hablan de carácter subjetivo o subjetividad de la experiencia. Por un lado, vimos que es común equiparar nociones como *para-mi-idad* (la experiencia es para-mí o se siente para mí) y *mi-idad* (la experiencia es mía o se siente mía) para referirse a la subjetividad. Pero hay autores que se resisten a tratar dichas nociones como sinónimos. Guillot (2016) y Sebastián (2012) han argumentado que son términos diferentes que capturan diferentes propiedades. Por ejemplo, Guillot sostiene que con el término "para-mi-idad" lo que se pretende capturar es que el sujeto se percata (además de objetos o cualidades) de la experiencia misma. Esto es, *del mismo estado*. Por otro lado, para el término "mi-idad", sostiene que el sujeto se percata de que *el estado es suyo*. En este caso, el sujeto no solo se percata del estado, sino de él mismo de alguna manera.

Para otros autores, esto no supone ningún problema. Por ejemplo, Kriegel (2009) ha argumentado que cuando el sujeto se percata de que la experiencia es para él (para-mi-idad), no se percata únicamente del estado, sino *de sí mismo* de alguna manera, ya que el sujeto mismo es parte del estado. Billon (2023), por su lado, concede que aunque sean términos diferentes, apuntan a la misma propiedad. En otras palabras, que son términos co-extensivos. Por ejemplo, sostiene que, si me percato de que la experiencia es *para-mí*, entonces me percato de que es una experiencia *mía* y, si me percato de que es mía, entonces no solo me percato de ella, sino *de mí misma* de alguna manera.

Si asumiéramos las visiones de Kriegel o Billon, entonces concederíamos que, para que haya carácter subjetivo, el sujeto no solo se percata de un estado, sino que se percata de él mismo. Sin embargo, esto por sí solo no resuelve la cuestión, pues también hemos visto que hay un debate respecto a cómo debe percatarse el sujeto de sí mismo para que haya

subjetividad de la experiencia. Esto se ejemplifica a partir de diferentes teorías de la consciencia.

Por un lado, está la teoría representacionista de orden superior según la cual cuando tenemos una experiencia, hay un estado de orden superior que representa "que *yo* estoy en un estado", en uno de orden inferior (Rosenthal, 1993, 1997; Gennaro, 1996, 2015). Por otro lado, está la teoría auto-representacionista según la cual hay un estado que se representa a sí mismo *como perteneciendo al sujeto*. Billon y Kriegel (2015) ejemplifican esta visión auto-representacionista de la siguiente manera: "todo estado consciente se representa a sí mismo (y más aún, se representa a sí mismo como perteneciendo al sujeto), y es en virtud de esta auto-representación que el sujeto se percata de él (como suyo)" [2015, p. 30].

No obstante, hay quienes no apoyan las teorías representacionistas. Algunos autores han afirmado que la percatación que tenemos de nuestros propios estados involucra una relación de familiaridad (*acquaintance*) con ellos (Levine, 2001). Otros afirman que esta familiaridad con los propios estados y con uno mismo no es más que una *percatación directa e inmediata* de nosotros mismos como sujetos de experiencias (Duncan, 2018; Billon, 2013; Williford, 2015). Donde, por "percatación directa e inmediata", se refieren a que la relación que tenemos con nosotros mismos no está mediada por ninguna entidad distinta (como una representación) o por un proceso distinto (un proceso reflexivo, por ejemplo).

Si bien hay diferentes visiones respecto a qué es el carácter subjetivo o cuál es la mejor manera de analizarlo, lo cierto es que, en todas ellas, el sujeto debe percatarse de sí mismo. Ya sea por medio de una representación (una auto-representación o una de orden superior) o por una relación de familiaridad. Y ya sea que al percatarse de sí mismo, también se percate de algún estado (que está en un estado de orden inferior o en uno del mismo orden).

No voy a comprometerme con una visión en particular. En lugar de eso, voy a mostrar que diferentes autores, además de estar de acuerdo en que para que haya carácter subjetivo, el sujeto debe percatarse de sí mismo, esta percatación debe cumplir con dos características constitutivas e íntimamente relacionadas que, si no cumple, no puede referir al carácter subjetivo o no califica como tal. A continuación presentaré tales características.

Pre-reflexividad: Sin importar qué visión se adopte, diversos autores (Duncan, 2018; Billon, 2013; Sebastián, 2020) han afirmado que la percatación de uno mismo es *pre-reflexiva*. Esto es, no necesitamos de ningún acto de reflexión, identificación o percepción de nosotros para percatarnos de que *nosotros mismos* tenemos una experiencia o estamos en un estado fenoménicamente consciente.

Por ejemplo, cuando yo tengo una experiencia de una manzana roja, me percato de que *yo veo* una manzana roja[14]. Para percatarme de que *soy yo* quien la ve, no necesito percibirme de alguna forma, tampoco necesito pensar o reflexionar sobre mí misma para percatarme de que yo tengo la experiencia, ni identificarme con una entidad en particular. Por ejemplo, identificarme con una entidad que resulta ser Andrea o con características físicas o de personalidad de dicha entidad. Siguiendo a Zahavi y Kriegel (2015), tampoco necesito identificarme como la dueña de dicha experiencia.

A partir de lo anterior, algunos autores (Gallagher, 2000; Duncan, 2018) sostienen que la percatación de uno mismo de manera pre-reflexiva o, también llamada *auto-consciencia pre-reflexiva*, es inmediata y no observacional, en tanto no necesitamos de un acto de consciencia perceptivo o reflexivo para percatarnos de nosotros mismos o en tanto no está mediada por ninguno de estos actos perceptivos o reflexivos.

No obstante, la idea de que me percato de mí cuando tengo una experiencia se ha puesto en duda a partir de la famosa objeción introspectiva de Hume. A continuación, voy a presentar dicha objeción y cómo, a partir de ella, se da lugar a otra característica que debe cumplir la subjetividad.

[14] Acabamos de ver que diversas teorías de la consciencia sostienen que, cuando el sujeto se percata de sí mismo, también se percata de alguna manera de un estado. Así, los filósofos de la consciencia conceden que, cuando un sujeto se percata de que *él ve X*, no solo se percata de él mismo y del objeto de la experiencia (X), sino de que él está en un estado visual de X o de que tiene una experiencia de X. Sin embargo, vimos que es motivo de debate en virtud de qué estado un sujeto se percata de que él está en cierto estado. Por ejemplo, un sujeto puede percatarse de que él ve X por medio de un estado de orden superior o por medio de una auto-representación. En la teorías de la familiaridad, se sostiene que la percatación del estado es directa (no implica una representación de por medio). Mi argumento no exige tomar una postura ante esta cuestión. Me estoy enfocando en la percatación del sujeto mismo, de manera que cuando el sujeto se percata de que él ve X, por ejemplo, se percata de que *él mismo es quien ve* o quien tiene una experiencia visual. Si bien se percata de alguna manera de un estado, esto es neutral respecto a lo anterior.

Elusividad: La objeción de Hume parte del hecho de que él no puede encontrarse a sí mismo en la introspección. "Por mi parte, cuando entro más íntimamente en lo que llamo *yo mismo*, siempre tropiezo con alguna percepción en particular, de calor o frío, luz o sombra, amor u odio, dolor o placer. Nunca puedo encontrarme en ningún momento sin una percepción, y nunca puedo observar nada más que la percepción" [Hume 1740/ 1888, p. 252].

A partir de esta objeción, podemos encontrar dos posturas. La postura escéptica según la cual, dado que no puedo encontrar un "yo" por introspección, el yo no existe. Esto explica que no pueda percatarme de mí misma cuando tengo una experiencia, sino solo de percepciones y/o de objetos de la percepción.

Por otro lado, está la postura no-escéptica. Esta concede que no podemos encontrar un "yo" por introspección, pero que esto no implica su inexistencia[15] o, de manera similar, que el hecho de que no podamos encontrar un "yo" por introspección no implica que no podamos percatarnos de nosotros mismos.

Para evitar el escepticismo, algunos autores han argumentado que la razón por la cual no encontramos un "yo" en la introspección es porque el yo no es un objeto que pueda introspectarse. Sin embargo, podemos percatarnos de nosotros mismos de distintas maneras a como nos percatamos de los objetos de la experiencia. Por ejemplo, Zahavi y Kriegel (2015) sostienen que la subjetividad no requiere que haya un *yo cualitativo* o un *yo objeto* aislado que uno pueda introspectar de la misma manera en la que uno introspecta el sabor de un helado o el aroma de un perfume. No obstante, todo objeto y/o cualidad de la experiencia se presenta al sujeto de la experiencia, *está dado* al sujeto y, de esta manera, hay algo que se siente *para-él* tener una experiencia o estar en un estado fenoménicamente consciente. Si hay algo que se siente *para- él* tener la experiencia, entonces hay un *yo* que la tiene o la disfruta, aun cuando el "yo" en cuestión no sea un objeto del que se percata. Luego, es elusivo a la introspección, pero aparece o figura en la experiencia.

[15] Debemos tener cuidado con esta afirmación. Las distintas maneras en las que se puede contestar a la objeción de Hume no se comprometen necesariamente con la existencia de un "yo" en el sentido en el que hay una entidad metafísica misteriosa aparte de un sujeto. Como lo plantearía Albahari (2006), como una entidad que podamos identificar como un "yo" que está "detrás" de un sujeto y que unifica todas las experiencias, pensamientos y acciones del sujeto. En cambio, puede haber una afirmación más débil en el sentido en el que si hay un yo que figura en la experiencia, este es idéntico al sujeto de la experiencia. Esto no implica postular una entidad misteriosa más allá del sujeto.

En una línea similar y retomando la característica de la pre-reflexividad, Sebastián (2020) ha argumentado que la auto-consciencia pre-reflexiva no supone la percatación de uno mismo como objeto o como una entidad en particular. De esta manera, sostiene que no hay una entidad que pueda encontrarse en la introspección, precisamente porque la auto-consciencia pre-reflexiva no es una cuestión de ser consciente de una entidad u objeto, así como podemos ser conscientes del objeto de la experiencia. No obstante, al igual que Zahavi y Kriegel, afirma que esto no implica que el sujeto no pueda percatarse de él mismo. Por ejemplo, sostiene que cuando él tiene la experiencia de una taza negra, su experiencia le informa o transmite el siguiente contenido[16]: "*yo veo* una taza negra". Luego, el *yo* es parte del contenido de la experiencia, aun cuando no sea un objeto en el sentido en el que lo es la taza negra.

Todo lo anterior no implica que de hecho no podamos percatarnos de nosotros mismos *como objeto*. Pensemos en casos en los que tenemos algunas experiencias de nuestros propios cuerpos. En dichos casos, somos el objeto de nuestra percepción. Sin embargo, lo que tratan de capturar los autores es que el "yo" al que se refieren, ese que está presente en toda experiencia posible, incluso en la de nuestros propios cuerpos, no es un objeto más o extra de la experiencia[17]. Para contrastarlo con casos en los que podemos ser objeto de la percepción, sostienen que uno se percata de sí mismo como sujeto. Así, el "yo" al que se refieren es un *yo-sujeto* o un *yo subjetivo*.

En resumen, los teóricos de la consciencia han debatido ampliamente cuál es la mejor manera de caracterizar la subjetividad de la experiencia. Sin embargo, pese a haber diferencias significativas entre ellos y/o entre sus teorías de la consciencia, he mostrado que,

[16] El autor caracteriza las experiencias en términos del contenido que un estado puede transmitir al sujeto de la experiencia. En este caso, su experiencia no le informa solo algo acerca del mundo (que hay una taza negra) sino algo de él mismo (que es él mismo quien ve). Los detalles de su propuesta no son relevantes para el propósito de este trabajo. Lo que me interesa mostrar es que, para el autor, el "yo" no aparece en la introspección (la elude) porque no es ningún objeto de la experiencia.

[17] Por ejemplo, si bien puedo percatarme de mis manos cuando *veo* que estoy escribiendo y en ese caso, soy objeto de mi percepción, también hay una dimensión subjetiva en el sentido en el que me percato de que *soy yo misma* quien está viendo mis manos. Podríamos decir que en este caso me percato de mí como sujeto y, a la vez, como objeto. Pero la percatación de mí misma como objeto o el "yo" como objeto ya no es parte de la dimensión subjetiva, pues hemos visto que la subjetividad refiere a la auto-consciencia pre-reflexiva y esta no requiere de la percepción de uno mismo. Más aún, cuando veo mis manos, puedo identificarlas como mis manos o fallar al identificarlas como mis manos. Luego, además de haber percepción, hay identificación y la pre-reflexividad tampoco requiere de ningún tipo de identificación.

cuando menos, están de acuerdo en que la subjetividad debe satisfacer las dos características mencionadas previamente.

Si la subjetividad de la experiencia debe satisfacer estas características para calificar como tal, una manera de analizar si acepta grados es aplicando los criterios de gradabilidad tanto a pre-reflexividad como a elusividad. Es decir, si acepta grados, depende de si las características necesarias que la constituyen pueden presentarse en grados[18].

1.5.2. Criterios de gradabilidad para la subjetividad de la experiencia

Hemos visto que los criterios de gradabilidad sirven para determinar si un adjetivo es gradable. Más aún, si el adjetivo es gradable, entonces denota una propiedad gradable. Si bien el término "elusividad" puede funcionar como adjetivo (refiere a lo elusivo o a lo que elude), el término "pre-reflexividad" puede funcionar tanto como adjetivo (refiere a lo pre-reflexivo) como adverbio de modo. Esto es, el sujeto *se percata* de él mismo *pre-reflexivamente* o de *manera pre-reflexiva*. Si precisamente a esto último nos referimos cuando decimos que la subjetividad es pre-reflexiva, podría pensarse que no tenemos los

[18] Podría objetarse que sólo estoy considerando dos características constitutivas de la subjetividad. A falta de una lista exhaustiva de características propias de ella, no puede concluirse definitivamente que la teoría gradualista no es viable. Sin embargo, dado lo controvertido que es el estudio de la subjetividad y, en general, el de la consciencia fenoménica, considero que será difícil encontrar otras características de las experiencias que no solo gocen de amplia aceptación dentro del estudio de la consciencia, sino que, además, sean claramente constitutivas de la subjetividad. Por ejemplo, Morales (2023) ha argumentado que *la fuerza mental* es una característica constitutiva de todas las experiencias en virtud de la cual éstas pueden ser más intensas o menos intensas. No obstante, él mismo reconoce que, en los estudios de la consciencia, es una característica que ha sido ignorada. Por otra parte, podemos cuestionar si es constitutiva de la subjetividad en sí y no de los aspectos cualitativos de los estados mentales (ya sean representacionales o no).
Otra característica que puede mencionarse en el contexto de los grados es *la confianza perceptiva*. Por ejemplo, Morrison (2016) ha argumentado que tenemos grados de confianza respecto a lo que percibimos. Un ejemplo que plantea es aquel en el que, si observamos un mantel a la luz de las velas, tenemos un alto grado de confianza de que es rojo en lugar de café, pero un grado bajo de confianza de que es carmesí en lugar de escarlata. Solo hasta que hay más luz, el grado de confianza de que es carmesí aumenta. Sin embargo, la propuesta de Morrison ha sido controversial en la filosofía de la percepción. Además, incluso si la aceptamos, no está del todo claro que los grados de confianza sean constitutivos de la subjetividad. Puedo tener más confianza de que lo que veo es carmesí y menos confianza de que lo que veo es escarlata, pero ¿puedo tener más confianza o menos confianza de que *soy yo* misma la que ve?
De nuevo, si bien no estoy dando una lista exhaustiva de características constitutivas de la subjetividad, al menos estas gozan de amplia aceptación. Por último, debemos considerar que mi objetivo no es refutar la teoría, sino plantear dificultades lo suficientemente serias que justifiquen la búsqueda de una alternativa teórica y, este objetivo puede cumplirse al considerarse las características que he mencionado previamente.

recursos para analizar si la pre-reflexividad se da en grados, pues, en primer lugar, "percatarse" es un verbo y, "de manera pre-reflexiva", ya vimos que funciona como adverbio.

No obstante, parece cuando menos intuitivo que podemos percatarnos de más o de menos y, en ese caso, la percatación o aquello de lo que nos percatamos podrían ser gradables. Por otro lado, parece que los criterios sí podrían funcionar para los adverbios. Por ejemplo, consideremos el adverbio "despacio" que expresa *la manera* en la que se realiza una acción. Podemos decir que "Juan camina más despacio que Andrés", así que cumple con el primer criterio de los adjetivos gradables. Por otro lado, tiene sentido preguntar "¿Qué tan despacio camina Juan?" y, por último, podemos contestar a esta pregunta diciendo que "Juan camina extremadamente despacio". De manera que el adverbio también cumple con los criterios dos y tres de los adjetivos gradables. Siendo así, todo parece apuntar a que sí tenemos los recursos para analizar también si la pre-reflexividad es gradable. Pasemos ahora al argumento.

Primero, voy a mostrar que si pensamos en una situación en la que un aspecto de la experiencia puede aumentarse o disminuirse, es porque nos enfocamos en aspectos de la experiencia diferentes de la subjetividad. Veamos el siguiente ejemplo.

Supongamos que mi amigo Axel y yo estamos en el museo del Prado admirando el cuadro *Las Meninas* de Velázquez. El cuadro y todos los elementos involucrados en él (las personas plasmadas en la pintura, los cuadros que están representados en el mismo cuadro, los detalles de los pisos, puertas, paredes, etc.) constituyen el objeto de la experiencia. Al ver el cuadro, tanto Axel como yo nos percatamos de él.

No obstante, si bien hay un sentido general en el que ambos nos percatamos del cuadro, es posible que Axel se percate más de él que yo, en tanto se percata de más características de él. Por ejemplo, Axel se percata de todas las personas que están plasmadas en la pintura y de todos los detalles que están alrededor en el cuadro. Por otro lado, yo solamente me percato de Velázquez y de *Las Meninas*. En este caso, podríamos decir que Axel se percata más del cuadro o del objeto de la experiencia y que yo me percato menos de él[19].

[19] Autores como Bayne (2010) y Lee (2022) han argumentado a favor de que los objetos y/o cualidades de los que nos percatamos se presentan en grados.

Si Axel se percata más del cuadro y yo menos, entonces se cumple el primer criterio de gradabilidad que refiere a construcciones comparativas. De manera más precisa, podríamos decir "Axel se percata más del cuadro de Velázquez que Andrea". Por otro lado, parece que tiene sentido preguntar "¿qué tanto se percata Axel del cuadro de Velázquez?" si consideramos que con esto nos estamos refiriendo a de qué tantas características del cuadro se percata. Como ya mencioné, se puede percatar de más características del cuadro o de todas ellas. Por último, dado que se percata de todas las características del cuadro, podríamos decir que se percata del cuadro *totalmente* o que se percata del cuadro *claramente*. Luego, también acepta modificadores de grados.

No obstante, en todos estos casos nos estamos refiriendo al objeto de la experiencia o a aspectos cualitativos de ella. Si hay un sentido en el que hay una diferencia de grados entre la experiencia de Axel y la mía, es en virtud del aspecto cualitativo. Pero, lo que necesitamos es que los grados se presenten en el aspecto subjetivo. Luego, a partir de este ejemplo, fallamos al encontrar grados en este último aspecto.

Ahora, voy a mostrar que si logramos enfocarnos en el aspecto subjetivo, las características constitutivas de dicho aspecto, a saber, pre-reflexividad y elusividad, no cumplen con los criterios de gradabilidad.

Pre-reflexividad y construcciones comparativas: Recordemos que, siguiendo a los teóricos de la consciencia, cuando nos percatamos del objeto de la experiencia, al mismo tiempo nos percatamos de nosotros mismos de manera pre-reflexiva o pre-reflexivamente. Una manera de plantear la construcción comparativa en relación con la pre-reflexividad y, apoyándonos en el ejemplo anterior, es la siguiente: "Axel es más pre-reflexivamente consciente cuando mira el cuadro *Las Meninas* que yo". Esto implicaría que yo soy menos pre-reflexivamente consciente que Axel. Sin embargo, esto no parece ser el caso. Partamos del supuesto de que la auto-consciencia pre-reflexiva es *inmediata*, en tanto no requiere de ningún acto de consciencia cognitivo o perceptivo que medie la percatación de uno mismo.

Para ser menos pre-reflexivamente consciente que Axel, la percatación de mí misma tendría que ser menos inmediata que la de él. Pero, si mi percatación fuera menos inmediata que la de él, entonces sería mediata. Es decir, tendría que estar mediada por algún acto cognitivo o perceptivo. En ese caso, la percatación de mí misma ya no sería pre-reflexiva. Y,

si no lo es, entonces desaparecería esta característica constitutiva de la subjetividad. Luego, la pre-reflexividad o la inmediatez no pueden aceptar grados hacia abajo.

¿Quizás podría aceptar grados hacia arriba, como sugiere la expresión "Axel es más pre-reflexivamente consciente"? De nuevo, ese no parece ser el caso. Si la auto-consciencia pre-reflexiva es inmediata, no parece tener sentido que la auto-consciencia pre-reflexiva de Axel sea más inmediata que la mía. Pues, ¿qué significaría esto? ¿Qué la suya es más no mediada? Esto no parece tener sentido. Si algo ya es inmediato, no puede ser más inmediato.

En suma, si yo soy pre-reflexivamente consciente o si me percato inmediatamente de mí, no puedo percatarme menos inmediatamente de mí que Axel porque, en ese caso, yo ya no sería pre-reflexivamente consciente. Y si Axel se percata pre-reflexivamente de él o inmediatamente de él, no puede percatarse más inmediatamente de él que yo, pues, en ambos casos, no hay ningún acto cognitivo ni perceptivo que esté mediando la percatación de cada uno de nosotros. En consecuencia, la pre-reflexividad/inmediatez no acepta construcciones comparativas.

Elusividad y construcciones comparativas: El razonamiento anterior es similar en el caso de elusividad. Si siguiendo a los teóricos de la consciencia, el yo siempre es elusivo a la introspección, no parece haber lugar para que sea más elusivo o menos elusivo. Por un parte, si al introspectar mi experiencia de *Las Meninas* no encuentro un yo en la introspección y, si al introspectar su experiencia de *Las Meninas* Axel tampoco encuentra un yo en su introspección, entonces el "yo" es igualmente elusivo para ambos. No parece que pueda haber un sentido en el que el "yo" de uno de nosotros sea más elusivo que el del otro.

Por otra parte, tampoco parece que el "yo" de alguno de nosotros pueda ser menos elusivo que el del otro. Pensemos que si el yo de alguno de nosotros fuera menos elusivo, entonces encontraría un rastro del yo en la introspección y, de ser así, ya no sería elusivo. Y, si no lo es, así como en el caso de la pre-reflexividad, también desaparecería esta característica constitutiva de la subjetividad. Para ver por qué, consideremos el siguiente ejemplo. Supongamos que, en lugar de estar observando el cuadro *Las Meninas*, me estoy observando a mí misma en el espejo. En este caso, puedo percatarme de que yo soy *la entidad* que estoy mirando en el espejo. Más aún, si me miro detenidamente en él, puedo percatarme de múltiples características físicas que poseo. Pero, al introspectar esta experiencia de mí

misma, *yo soy el objeto* de la experiencia y, al ser el objeto de ella, encontraría un "yo" por introspección[20]. Sin embargo, hemos visto que el "yo" que siempre escapa o elude la introspección es un *yo-sujeto* que no necesita identificarse con ninguna entidad o con características de dicha entidad. Luego, o encuentro un yo en la introspección, pero este no es el yo-sujeto que postulan los teóricos de la consciencia o no encuentro nunca un yo en la introspección, pero esto implica que no hay menos elusividad o que el yo pueda ser menos elusivo.

Pre-reflexividad y sentido a las preguntas del tipo "¿qué tan alto es X?" o "¿qué tanta altura tiene X?": Para este criterio podemos plantear la pregunta de la siguiente manera: "¿Qué tanto se percata S (Axel o Andrea, por ejemplo) de sí mismo pre-reflexivamente?" Si la pregunta tiene sentido, depende de qué puede significar "qué tanto se percata S de X". Como sugerí anteriormente, al preguntar "¿Qué tanto se percata Axel del cuadro de Velázquez?", puede significar de qué tantas características o cualidades se percata Axel al ver el cuadro. Como vimos, puede percatarse de más o de todas ellas. En este caso, la pregunta tiene sentido. Sin embargo, nos estamos refiriendo al objeto de la experiencia, no al sujeto.

No obstante, parece que también tiene sentido preguntar "¿qué tanto se percata uno de sí mismo?" cuando se ve detenidamente en el espejo, por ejemplo, o cuando piensa en el tipo de persona que fue o que es en la actualidad. Pero, de cualquier manera, este tipo de pregunta es similar a la anterior en tanto nos estamos refiriendo a características (ya sean físicas o de personalidad) del sujeto *como objeto*. Es decir, el sujeto se percata de sí mismo como objeto o es el objeto de una percepción o pensamiento. Sin embargo, hemos visto que la pre-reflexividad está libre de actos perceptivos o cognitivos. Esto es, no está mediada por ninguno de estos actos y/o la auto-consciencia pre-reflexiva no requiere percatarse de uno mismo como objeto ni de características o cualidades de uno mismo.

Siendo así, no parece tener sentido preguntar "¿qué tanto se percata el sujeto de sí mismo *pre-reflexivamente*?" pues ya vimos que la construcción "qué tanto" refiere,

[20] Cabe señalar que, incluso si puedo percatarme de mí como objeto y no escapar a la introspección, también en este tipo de casos me percato de mí como sujeto. Por ejemplo, me percato de que *yo veo* cierta entidad en el espejo (independientemente de si me identifico con dicha entidad o con las características de ella) y este aspecto, este yo-sujeto, es el que sigue permaneciendo elusivo a la introspección.

precisamente, a de qué tantas características o cualidades se percata el sujeto. Luego, la pregunta no parece aplicar a pre-reflexividad.

Elusividad y sentido a las preguntas del tipo "¿qué tan alto es X?" o "¿qué tanta altura tiene X?": En este caso, podemos revisar el tipo de respuestas que se pueden dar a las preguntas como "¿Qué tan alto es X?". Podríamos contestar "X es muy alto" (o *extremadamente* alto), "X no es tan alto" (o es *ligeramente* alto) o "X es tan alto como Y". Este tipo de respuestas parecen apuntar a la medida en la que se presenta una propiedad. Se puede presentar en gran medida (más), en menor medida (menos) o en la misma medida. Sin embargo, si se presenta en la misma medida, no parece haber grados.

Pasemos ahora a elusividad. Ya vimos con el criterio de las construcciones comparativas que el "yo" no puede ser más elusivo ni menos elusivo a la introspección. Luego, no tendría sentido preguntar algo como "¿qué tan elusivo a la introspección es el yo de algún sujeto (el de Axel, por ejemplo)?" pues, si no puede ser más elusivo o menos elusivo, no podríamos contestar "el yo de Axel es *muy* elusivo o *ligeramente* elusivo a la introspección". Esto es, la elusividad (del yo) no se da ni en gran medida ni en menor medida.

Por supuesto, en el lenguaje natural podríamos decir, por ejemplo, que el zorro es *muy* elusivo o que el lince es *menos* elusivo que el zorro. En estos casos, la elusividad se puede presentar en mayor o menor medida. Sin embargo, debemos considerar que, en estos casos, nos estamos refiriendo a entidades específicas y, hemos visto a lo largo de esta sección, que el yo que elude a la introspección no es un objeto o entidad específica que se pueda introspectar como otros objetos de la experiencia. De manera similar, incluso si el yo como objeto puede ser más elusivo o menos elusivo, el yo que elude siempre a la introspección es un yo-sujeto.

Ahora, si cuando dos sujetos introspectan cualquier experiencia y ninguno encuentra un yo en la introspección, entonces ya vimos que esto implica que el yo es igualmente elusivo para ambos. Siendo así, si preguntáramos "¿qué tan elusivo a la introspección es el yo de Axel?", solo podríamos contestar "es tan elusivo como el de Andrea" y, en este caso, la respuesta no sugiere grados.

Así, ni la pre-reflexividad ni la elusividad tampoco cumplen con el segundo criterio de los adjetivos gradables. Ahora bien, el tercer criterio dicta que los adjetivos gradables deben aceptar modificadores de grados. No obstante, si estas características no cumplen con los primeros dos criterios, entonces no son gradables y, si no lo son, no tendría sentido aplicar modificadores de grados a características no gradables. Esto puede reflejarse, por ejemplo, con el caso de elusividad. Como vimos, no podríamos decir que el yo de algún sujeto es *extremadamente* elusivo o *ligeramente* elusivo, pues el yo no puede ser más elusivo ni menos elusivo a la introspección.

En resumen, en la primera parte del argumento señalé que, cuando logramos identificar grados en las experiencias, estos corresponden al objeto de la experiencia o al aspecto cualitativo de ella. En la segunda parte, señalé que, al centrarnos en la subjetividad, sus características constitutivas no cumplen con los criterios de gradabilidad. Sobre todo, si nos centramos en las construcciones comparativas, ni la pre-reflexividad ni la elusividad pueden aceptar grados hacia abajo debido a que estas desaparecerían. Y, si son constitutivas de la subjetividad, al desparecer estas, ya no habría subjetividad. Luego, la teoría también presenta dificultades teóricas.

En la última sección mostraré que, pese a que la teoría gradualista no considera evidencia empírica para dar cuenta de la anomalía de las experiencias de los pacientes, hay una serie de estudios que son difíciles de reconciliar con lo que plantea la teoría.

1.6. Dificultades con la evidencia empírica

Si bien es cierto que, como acabo de señalar, la teoría no considera evidencia empírica en su estudio, esto no significa que no tengamos evidencia neurológica disponible que, en principio, plantea problemas para la teoría.

Diferentes estudios de neuroimagen han mostrado que los pacientes despersonalizados exhiben una activación reducida o nula actividad en áreas cerebrales involucradas en el sistema límbico, así como en áreas íntimamente conectadas con dicho sistema. Por ejemplo, Sierra y David (2011) mencionan que tanto a un grupo de pacientes despersonalizados, como a un grupo de sujetos sanos y a otro grupo con trastorno obsesivo

compulsivo, se les pidió observar una serie de imágenes aversivas y neutras mientras se les realizaba una resonancia magnética funcional.

El estudio mostró que, a diferencia de los sujetos sanos y aquellos con trastorno obsesivo compulsivo, "los pacientes con despersonalización (…) afirmaron que, si bien podían comprender el contenido de las imágenes, no experimentaron ninguna respuesta emocional subjetiva" [2011, p. 104]. Por otra parte, mientras que en los sujetos sanos y en los pacientes obsesivos hubo una activación en la ínsula anterior en respuesta a las imágenes aversivas, se encontró que en los pacientes despersonalizados no hubo activación de esta área cuando vieron dichas imágenes.

En otros estudios, Lemche et al. (2007, 2008) realizaron también una resonancia magnética funcional a pacientes con despersonalización y a un grupo de sujetos sanos mientras observaban una serie de imágenes de caras tristes y felices. Mientras que en el grupo de control se observó un aumento en la actividad de la amígdala y el hipotálamo cuando la expresión emocional aumentaba (las caras se hacían más felices o más tristes), los pacientes despersonalizados exhibieron un patrón opuesto. Es decir, la actividad en estas áreas era menor cada que aumentaba la intensidad de la expresión emocional.

Ahora bien, como sugieren estos estudios, el sistema límbico y/o las áreas que lo componen (entre ellas la amígdala y el hipotálamo) juegan un papel importante en la generación de emociones o respuestas emocionales, al igual que las áreas que tienen una conexión con este sistema, como la ínsula anterior. Por ejemplo, Sierra y David (2011) han señalado que "investigaciones recientes han identificado la ínsula anterior como una región cortical crucial para la experiencia de las emociones (Critchley, 2005; Morris, 2002)" [Sierra y David, 2011, p. 104].

Si tomamos en cuenta estos estudios, entonces la evidencia no es convergente con la teoría. Lo que los estudios mostraron fueron déficits específicos, a saber, emocionales o afectivos que, a su vez, derivan de una actividad anómala en las áreas que subyacen a este tipo de estados. Sin embargo, la teoría no postula déficits afectivos ni siquiera para dar cuenta de los reportes que refieren a emociones. Lo que postula es una anomalía estructural de la consciencia derivada de una atenuación de la subjetividad.

Para que la evidencia sea convergente con la teoría, esta podría afirmar que la subjetividad está relacionada con la afectividad, de manera que, si hay déficits afectivos, entonces hay déficits subjetivos o, alternativamente, que el déficit subjetivo *es* un déficit afectivo. Esto implicaría que la subjetividad es afectiva por naturaleza.

No obstante, ninguna de las opciones es viable para la teoría. Al postular una atenuación de la subjetividad, incluso para los estados afectivos, el gradualista busca dar cuenta de ellos en lo que llama "términos más fundamentales" (véase Billon, p. 42). Es decir, el déficit subjetivo es más fundamental que el afectivo y/o hay déficits afectivos porque hay déficits subjetivos y no al revés. Respecto a la segunda opción, el gradualista no está convencido de que la subjetividad sea afectiva o de tipo afectivo, pues sostiene lo siguiente: "esta es una tesis muy fuerte y no estoy seguro de cómo se podría argumentar convincentemente a favor de ella" [Billon, 2016, p. 43].

En vista de lo anterior, podemos concluir que la evidencia empírica no solo no respalda la teoría gradualista, sino que, al parecer, esta última no tiene los recursos para incorporarla en su marco gradualista.

1.7. Conclusión

La teoría gradualista interpreta las experiencias anómalas de los pacientes despersonalizados en términos de una atenuación del carácter subjetivo de la experiencia. Siendo así, no solo ofrece una interpretación de dónde proviene la anomalía de las experiencias de estos pacientes, sino que resguarda el principio subjetivo según el cual toda experiencia es necesariamente subjetiva. Esto es, si la subjetividad está atenuada, en lugar de ausente, entonces las experiencias de estos pacientes siguen siendo subjetivas, aunque en un grado menor. Sin embargo, he argumentado que la teoría enfrenta tres dificultades diferentes para dar cuenta de esta anomalía.

Primero, presenta problemas interpretativos. La teoría sugiere que debemos tomar los reportes de los pacientes literalmente y predice que, si un sujeto no siente claramente que cierto estado mental es suyo, entonces éste expresará que es como si no fuera suyo o como si no le perteneciera. No obstante, a partir de esta sugerencia y de su predicción, vimos que, la expresión "como si" es ambigua. Esto nos llevó a buscar otros criterios lingüísticos que

pudieran indicar grados de subjetividad. A partir de dichos criterios, vimos que la dificultad predominante es que muchos sujetos no utilizan la expresión "como si X no fuera mío" en sus reportes. La segunda dificultad, es que encontramos expresiones gradables que remiten a la afectividad y no a la subjetividad. La tercera, es que muchos reportes no sugieren grados de ningún tipo y, la cuarta, que algunos reportes no se relacionan directamente con estados mentales.

Ante estas dificultades, argumenté que la teoría tenía que ser enriquecida con hipótesis auxiliares. Sin embargo, estas resultaron problemáticas por ser poco plausibles, poco específicas o porque no se ofrecen razones independientes para adoptarlas.

Segundo, presenta problemas teóricos. Mostré que, pese a haber diferentes concepciones de la subjetividad, suele aceptarse que esta debe cumplir con dos características constitutivas para calificar como tal: pre-reflexividad y elusividad. Al utilizar los criterios de gradabilidad como una herramienta heurística, argumenté que, cuando encontramos grados en las experiencias, estos corresponden a aspectos no-subjetivos. Sin embargo, cuando logramos centrarnos en el aspecto subjetivo, las características constitutivas de la subjetividad no satisfacen los criterios de gradabilidad pues, en términos generales, desaparecerían dichas características. Y, al desaparecer, ya no habría subjetividad.

Tercero, presenta problemas con la evidencia empírica. A partir de los estudios de neuroimagen disponibles, vimos que los pacientes despersonalizados exhiben actividad anómala en áreas que juegan un papel importante en la generación de emociones o respuestas emocionales, lo cual apuntaría a déficits afectivos. No obstante, vimos que la teoría no postula déficits afectivos. Luego, la evidencia no es convergente con la teoría. Plantee algunas opciones que tendría la teoría para converger con la evidencia. Sin embargo, mostré que ninguna de ellas era viable y, de ser así, al parecer no tiene los recursos para incorporar la evidencia en el marco gradualista.

Si bien mi objetivo no ha sido refutar la teoría, las tres dificultades previamente mencionadas muestran que dicha teoría enfrenta serios problemas para dar cuenta de la anomalía de las experiencias de estos pacientes.

Mientras los partidarios de la teoría gradualista buscan soluciones a estas dificultades, tenemos razones suficientes para buscar una teoría alternativa que quizás nos ahorre el esfuerzo de enfrentar tales problemas.

Capítulo 2

La teoría negacionista

2.1. Introducción

En el primer capítulo, vimos que una experiencia es un estado fenoménicamente consciente y que este tipo de estados difieren en virtud de su *carácter cualitativo* (por ejemplo, por los objetos y/o cualidades de los que nos percatamos), mientras que tienen en común que estos siempre se presentan como tuyos o para-ti. Es decir, que comparten la característica de la *subjetividad*. De la misma manera, vimos que si las experiencias o estados fenoménicamente conscientes siempre son para-ti o son tuyos, entonces la subjetividad es constitutiva de toda experiencia posible, de manera que no puede haber experiencias sin subjetividad. Esta observación quedaba capturada por el llamado *principio subjetivo* según el cual "necesariamente, un estado mental M exhibe consciencia fenoménica solo si M exhibe subjetividad" [Billon y Kriegel, 2015, p. 30].

Asimismo, vimos que, dada la naturaleza de los reportes de los pacientes despersonalizados, estos podrían mostrar que el principio subjetivo es falso. Al igual que la teoría gradualista, la negacionista busca dar cuenta de la anomalía de los estados de estos pacientes sin poner en riesgo el principio subjetivo.

En términos generales, la teoría plantea que, efectivamente, los estados que reportan estos pacientes no son subjetivos, pero, dado el principio subjetivo, si no lo son, tampoco son fenoménicamente conscientes. Debido a que los pacientes pueden reportar estos estados, la teoría plantea que tienen acceso a dichos estados, que no son fenoménicamente conscientes.

En este capítulo, voy a mostrar que, de manera similar a la teoría gradualista, la negacionista enfrenta diferentes tipos de dificultades para dar cuenta de la anomalía de los estados de estos pacientes. En primer lugar, enfrenta problemas interpretativos. En segundo lugar, enfrenta problemas explicativos y, por último, también enfrenta problemas con la evidencia empírica disponible.

Antes de presentar dichos problemas, en la sección 2.2, voy a exponer en qué consiste la teoría negacionista defendida por Billon y Kriegel (2015), así como las predicciones que se desprenden de ella. De manera general, la teoría arroja dos posibles predicciones: si un

estado no es subjetivo y, por tanto, es fenoménicamente inconsciente, pero hay acceso a dicho estado, el sujeto expresará que dicho estado no se siente como si fuera suyo o, alternativamente, si hay un estado que no es fenoménicamente consciente, pero el sujeto tiene acceso a dicho estado, entonces expresará que es consciente e inconsciente.

En la sección 2.3, voy a mostrar que la teoría negacionista enfrenta diferentes dificultades interpretativas. Esto se debe a que, de manera general, muchos de los reportes de los pacientes no confirman las predicciones de la teoría. Como vimos en el primer capítulo, son muy pocos los sujetos que expresan que X es como si no fuera suyo. Asimismo, son muy pocos los sujetos que expresan que son conscientes e inconscientes. Relacionado con lo anterior, también hay discrepancias entre los reportes de los pacientes y las predicciones de la teoría. Por un lado, aunque algunos sujetos expresan consciencia e inconsciencia, no es claro que se refieran inequívocamente a los sentidos de consciencia que plantea la teoría. Por otro lado, una gama amplia de reportes no sugieren que los estados de los pacientes no sean subjetivos y, por tanto, no tenemos buenas razones para concluir que no son fenoménicamente conscientes. Por último, la teoría no especifica cómo se puede acceder a la información de un estado sin recurrir a experiencias o estados fenoménicamente conscientes. Ante estas dificultades, voy a argumentar que, de manera análoga a la teoría gradualista, la negacionista tendría que enriquecerse con hipótesis auxiliares para que los reportes confirmen las predicciones que plantea. Sin embargo, también mostraré que las hipótesis más naturales que se podrían añadir, resultan problemáticas.

Si bien el último problema que mencioné anteriormente se da a nivel interpretativo, en la sección 2.4, voy a señalar que éste da lugar a un problema más general, a saber, que la teoría *asume* que hay acceso *sin explicar* de qué manera se podría acceder a estados que, supuestamente, carecen de fenomenología propia. Para tal fin, en la sub-sección 2.4.1, voy a recurrir al caso hipotético de la súper visión ciega que plantea Block (1995/2007) para ilustrar la posibilidad conceptual de acceso sin fenomenología. A partir de este caso, en la sub-sección 2.4.2, voy a argumentar que, si los defensores de la teoría negacionista están tomando las distinciones de acceso y fenomenología de Block, entonces deberíamos de esperar que los casos de despersonalización se expliquen de manera similar al caso hipotético de la súper visión ciega. En particular, la teoría debería postular un mecanismo de acceso análogo al que

postula Block en su caso hipotético. Sin embargo, esto no es lo que encontramos y, al no postular dicho mecanismo, la teoría no ofrece una concepción del acceso que no explote la fenomenología del estado al que se accede.

En la sección 2.5, voy a mostrar que la teoría negacionista también enfrenta dificultades con la evidencia empírica disponible. Por un lado, la teoría plantea que los pacientes despersonalizados no tienen que hacer ningún esfuerzo para acceder a sus estados fenoménicamente inconscientes. Lo que sugiere que, en general, no hay problemas con el acceso. Por otro lado, si los estados de los pacientes presuntamente carecen de fenomenología, deberíamos encontrar daños en áreas que pueden contribuir a la fenomenología de algunos estados. En la sub-sección 2.5.1, voy a presentar las áreas cerebrales que se han asociado con el acceso, así como pruebas neurológicas que sugieren que hay dificultades con el acceso. En la sub-sección 2.5.2, voy a señalar que, si bien es cierto que en la actualidad no tenemos una idea clara respecto a las áreas cerebrales que subyacen a la fenomenología o a la consciencia fenoménica, podríamos asumir que ciertas áreas involucradas en estados individuales (percepciones o emociones, por ejemplo) contribuyen a la fenomenología de dichos estados. No obstante, voy a mostrar que, aun asumiendo lo anterior, la evidencia no parece converger con lo que plantea la teoría.

Por último, en la sección 2.6, voy a presentar las conclusiones del capítulo. De la misma manera que con la teoría gradualista, mi objetivo no es refutar la teoría negacionista, sino mostrar que las dificultades mencionadas son lo suficientemente serias para justificar la búsqueda de una teoría alternativa.

2.2. La teoría negacionista

De acuerdo a la teoría negacionista, el problema con los pacientes despersonalizados es que *algunos* de sus estados no son subjetivos (véase Billon y Kriegel, pp. 36 y 44-45). Sin embargo, contrario a lo que podría pensarse, esto no significa que haya experiencias o estados fenoménicamente conscientes sin subjetividad. Debido a que aceptan el principio subjetivo, los negacionistas plantean que, si los estados de los pacientes no son subjetivos, entonces tampoco son fenoménicamente conscientes. Si no son fenoménicamente conscientes, entonces no hay nada que se sienta para ellos estar en dichos estados.

Los estados en cuestión pueden ser de un solo tipo, como pensamientos, sensaciones corporales, emociones o estados perceptivos. En los casos extremos, cuando la despersonalización ha alcanzado su clímax, los negacionistas afirman que *la mayoría* de los estados dejan de ser fenoménicamente conscientes. A los estados que no son fenoménicamente conscientes los denominan *estados alienados*. Es decir, estados que no se sienten como si fueran propios o que se sienten como si fueran ajenos. Así, el estado alienado puede ser una percepción, un pensamiento, sensaciones corporales o, como sugieren los negacionistas, múltiples estados mentales.

Por otra parte, afirman que debido a que los pacientes son capaces de notar y reportar estos estados (de otra manera, sostienen que no los molestarían), tienen acceso a ellos. En otras palabras, tienen consciencia de acceso a algunos estados que no son fenoménicamente conscientes o que no tienen fenomenología propia. Podemos pensar que la anomalía proviene de estos dos factores, pues parece ser atípico que un sujeto pueda reportar estados que no son fenoménicamente conscientes.

Sabemos ya qué es la consciencia fenoménica o qué es un estado fenoménicamente consciente, más no qué es la consciencia de acceso. Para entender qué es, podemos apelar a la noción paradigmática de consciencia de acceso de Block (1995/2007). Para el autor, un estado posee consciencia de acceso o es A-consciente si el contenido de ese estado (o la información que porta) está disponible para el razonamiento, el control racional de la acción y el habla[21]. A partir de esta caracterización, Block sostiene que la función de la consciencia de acceso es "permitir de alguna manera que la información representada en el cerebro se utilice para razonar, informar y guiar racionalmente la acción" [Block, 1995/2007, p. 160]. En otras palabras, su función es procesar la información representada y ponerla a disposición de otros sistemas, como aquellos encargados del reporte, la formación de creencias y de la acción (Sebastián, 2022).

[21] Por *control racional*, Block pretende distinguir entre casos en los que los sujetos actúan o reportan algo por inclinación o por incitación de alguien más y casos en los que actúan o reportan porque tienen una razón para hacerlo y/o porque lo hacen de manera voluntaria o espontánea. Solo en estos últimos casos hay control racional. En la sub-sección 2.4.1 desarrollaré más detalladamente la noción de acceso, así como las características que lo definen.

Así, la idea de que algunos estados de los pacientes despersonalizados son conscientes en el sentido de acceso, pero no en el sentido fenoménico, queda capturada de la siguiente manera:

> Considere que los reportes de los pacientes indican que se percatan de sus estados alienados de manera similar a como nosotros nos percatamos de nuestros estados fenoménicamente conscientes. En particular, no tienen que hacer ningún esfuerzo para acceder a sus estados alienados (…) podríamos decir que el acceso de los pacientes a sus estados alienados parece 'inmediato'. Presuntamente, hay un sentido de 'consciencia' según el cual cualquier estado mental al que se tiene acceso aparentemente inmediato es, por definición, consciente. Podríamos llamarlo el sentido *reflexivo* de 'consciencia'. Se podría argumentar entonces que la consciencia reflexiva (…) no implica a priori la consciencia fenoménica (…) y que aunque los estados alienados son conscientes en el sentido reflexivo, no son conscientes en el sentido fenoménico[22] [Billon y Kriegel, 2015, p. 44].

A partir de todo lo anterior, la teoría plantea una conjunción: un estado específico o múltiples estados que no son subjetivos, no son fenoménicamente conscientes y tales estados son accesibles al sujeto (o tiene acceso a ellos). O, un estado o múltiples estados son fenoménicamente inconscientes y son A-conscientes o accesiblemente conscientes. Por ejemplo, si el estado alienado es un pensamiento y si éste es accesiblemente consciente, el sujeto puede reportar, en primer lugar, que no lo siente como si fuera suyo o que lo siente como si fuera ajeno y, en segundo lugar, dada la caracterización de consciencia de acceso de Block, puede reportar la información que transmite el pensamiento alienado.

Ahora bien, al igual que la teoría gradualista, la negacionista también sugiere que debemos tomar los reportes de los pacientes de manera literal. Esto es, si el sujeto dice que cierto estado no se siente como si fuera suyo, entonces literalmente no lo siente propio. Luego, no es subjetivo ni fenoménicamente consciente. Los negacionistas sostienen que debemos suponer lo anterior debido a que "el contenido de los reportes es que los estados

[22] Podría pensarse que el sentido reflexivo de consciencia no es equivalente al sentido de acceso que plantea Block. Sin embargo, tenemos razones para pensar que, de hecho, los negacionistas equiparan ambos, pues sostienen lo siguiente: "Usando las distinciones de Block (1995), podríamos decir que un estado es consciente en el sentido de consciencia reflexiva cuando su sujeto es A-consciente de él, o cuando es reflexivamente-consciente" [Billon y Kriegel, 2015, p. 50].

relevantes no se sienten como si le pertenecieran a uno, y lo que llamamos subjetividad es precisamente esa sensación" [Billon y Kriegel, 2015, p. 36].

Sin embargo, a diferencia de lo que sugiere la teoría gradualista, la teoría negacionista no toma la expresión "como si" como un marcador implícito de grados de subjetividad. En cambio, la entienden como una expresión que indica que el sujeto no es delirante: "los pacientes despersonalizados aún *creen* que los estados alienados son suyos. Los pacientes dicen que sienten *como si* esos estados no les pertenecieran. Hay una sensación de alienación, pero los pacientes no *respaldan* la sensación. Esto es, no toman la apariencia de alienación al pie de la letra y, por tanto, se auto-atribuyen esos estados. En otras palabras, no son delirantes" [Ibíd., p. 35]. Siguiendo la sugerencia de tomar los reportes literalmente, también indican que algunos pacientes específicamente expresan que son conscientes (en el sentido de acceso) e inconscientes (en el sentido fenoménico).

A partir de todo lo anterior, la teoría negacionista arroja dos predicciones generales: 1) si un estado específico o múltiples estados no son subjetivos y, por tanto, son fenoménicamente inconscientes y, si además los pacientes tienen acceso a dichos estados, entonces estos expresarán que no se sienten como si fueran suyos, pero informarán el contenido de los estados o la información que portan, o 2) si un estado o múltiples estados son fenoménicamente inconscientes, pero accesibles, entonces expresarán que son conscientes en un sentido (de acceso), pero no en otro (fenoménico).

No obstante, de seguir la sugerencia de tomar los reportes de manera literal, surgiría una serie de dificultades interpretativas. En la siguiente sección presentaré estas dificultades.

2.3. Dificultades interpretativas: la interpretación negacionista de los reportes

En un sentido general, la teoría negacionista no puede dar cuenta de un gran número de reportes pues, como vimos en el primer capítulo, son pocos los reportes donde los pacientes expresan que X no se siente como si fuera suyo o como si no le perteneciera y, como también ya vimos, estos refieren a partes corporales, no a estados mentales[23]. De la misma manera,

[23] Cabe señalar que, en estos casos, los negacionistas también afirman que para la negación de partes corporales y movimientos corporales automáticos, los estados relevantes son las sensaciones corporales y las intenciones para la acción. Por motivos de extensión no voy a considerar estos casos para esta teoría. Sin embargo, podemos pensar que se enfrenta a los mismos problemas que la gradualista. En el caso de las sensaciones corporales, los negacionistas dirían que las sensaciones no son subjetivas, pero ya vimos que el

son pocos los reportes donde los pacientes expresan que son conscientes en un sentido, pero no en otro.

Relacionado con estos problemas generales, también hay discrepancias entre las predicciones de la teoría y los reportes de los pacientes. En primer lugar, aunque algunos de ellos reportan que son conscientes en un sentido y no en otro, no es claro que se refieran inequívocamente a los sentidos de consciencia a los que apela la teoría. En segundo lugar, por el lado de la fenomenología, muchos de los reportes de los pacientes no sugieren que sus estados no sean subjetivos y, por tanto, que no sean fenoménicamente conscientes. Por último, aunque este problema lo voy a desarrollar a detalle en la siguiente sección, voy a señalar que, de manera general, la teoría no especifica cómo se puede acceder a la información de un estado sin recurrir a experiencias o estados fenoménicamente conscientes.

En vista de estos problemas, voy a argumentar que, para que los reportes cumplan con las predicciones de la teoría, ésta también necesita ser enriquecida con hipótesis auxiliares que, a su vez, resultan ser problemáticas debido a que son poco plausibles o porque parecen ir en contravía de los mismos reportes o de lo que plantea la teoría.

Voy a proceder de la misma manera que con la teoría gradualista y sus dificultades interpretativas. Comenzaré por plantear el problema específico, voy a citar reportes que lo ejemplifiquen, los supuestos o hipótesis que se podrían añadir para solucionarlos y voy a señalar los problemas con dichos supuestos[24].

Consciencia e inconsciencia: Como hemos visto, la teoría predice que si un estado o múltiples estados son fenoménicamente inconscientes, pero accesibles (tiene acceso a ellos), entonces los pacientes podrían reportar que de alguna manera son conscientes (en el sentido de acceso) e inconscientes (en el sentido fenoménico). Pero, ¿tenemos razones para pensar que sus reportes apuntan en esta dirección?

---

hecho de que las sensaciones no sean subjetivas (en absoluto o en menor grado) no necesariamente está relacionado con que se niegue la parte corporal o que las sensaciones corporales no fundamentan la pertenencia corporal. Respecto a las intenciones para la acción, también podemos cuestionar por qué es la intención y no otro aspecto de la experiencia la que no es fenoménicamente consciente.

[24] Al igual que con la teoría gradualista, voy a citar reportes que la misma teoría negacionista interpreta en términos de ausencia de fenomenología y presencia de acceso. De la misma manera, no estoy considerando que no exista ninguna hipótesis auxiliar que pueda ayudar a dar cuenta de los reportes. Pero, al menos las más naturales pueden resultar problemáticas.

Uno de los pacientes reporta lo siguiente: “simplemente me hundo en una especie de inconsciencia. Soy lo suficientemente consciente para saber que suceden cosas a mi alrededor, pero nada parece registrarse” [Shorvon, 1946, p. 784]. Lo primero que podríamos preguntarnos es cuál o cuáles son los estados que no son fenoménicamente conscientes pues, en sí mismo, el reporte no lo específica. Sin embargo, señalaba anteriormente que los negacionistas afirman que, cuando la despersonalización alcanza su clímax, la mayoría de los estados dejan de ser fenoménicamente conscientes y añaden que entonces “los sujetos pueden describirse a sí mismos como totalmente inconscientes, incluso como zombies” [Billon y Kriegel, 2015, p. 35]. Concedamos que este reporte ejemplifica este caso.

Ahora, la teoría parece suponer, en primer lugar, que el sujeto se describe como “suficientemente consciente” porque sabe que suceden cosas a su alrededor, a la vez, porque tiene acceso a la información de sus estados; pero, en segundo lugar, se describe como inconsciente o como “hundiéndose en una especie de inconsciencia” porque, aunque lo sabe, “nada parece registrarse”, suponiendo, al parecer, que esto último implica que *no hay nada que se sienta para el sujeto ver lo que sucede a su alrededor*, por ejemplo o, que el sujeto se refiere precisamente a esto cuando afirma que nada parece registrarse.

No obstante, hay problemas con ambos supuestos. Por una parte, en la sub-sección 1.4.2 del primer capítulo, mencioné que es implausible suponer que cuando un sujeto se describe como muerto quiera decir que se siente *fenoménicamente muerto*, debido a que es poco probable que la gente no especializada en filosofía conozca la diferencia entre tener estados fenoménicamente conscientes e inconscientes. Ahora, este caso es similar respecto a este punto; resulta implausible suponer que el sujeto se describe como suficientemente consciente en el sentido de acceso y como inconsciente en el sentido fenoménico, porque es poco probable que conozca esta distinción. En otras palabras, difícilmente podría trazar distinciones finas de filosofía de la mente que solo fueron acuñadas a finales del siglo XX y que sólo están disponibles para un reducido número de expertos.

Aunado a este problema, suponer lo anterior contradice la máxima metodológica de los negacionistas según la cual deberíamos tomar los reportes de los pacientes *literalmente*. De nuevo, si es poco probable que los pacientes conozcan la distinción entre conciencia fenoménica y consciencia de acceso, es difícil pensar que literalmente se refieren a esto

cuando se describen como conscientes e inconscientes. Es decir, que se refieren a que son suficientemente conscientes en el sentido de acceso e inconscientes porque nada parece registrarse y, además, que esto último literalmente significa que no hay nada que se siente para ellos ver lo que sucede a su alrededor o que no hay nada que se siente para ellos estar en diferentes estados.

Por otra parte, teóricos sofisticados como Billon y Kriegel saben muy bien que la fenomenología no se define por el registro de información, sino por cómo se siente para un sujeto estar en un estado o tener una experiencia. Así, no tienen buenas razones para concluir la ausencia de fenomenología a partir del hecho de que "nada parece registrarse". Por "registro" podemos entender procesamiento y/o recopilación de información y almacenamiento de ella y estos son procesos que teóricos sofisticados como Billon y Kriegel deberían identificar como parte de la consciencia de acceso, no de la consciencia fenoménica[25].

Estos mismos problemas los podemos encontrar con el siguiente reporte: "Soy como un zombie, incapaz de asimilar ninguna información" [Sierra, 2009, p. 51]. De nuevo, es poco probable que un sujeto sepa qué es un zombie en el sentido que los filósofos de la mente plantean, precisamente como una criatura que no tiene estados fenoménicamente conscientes, pero sí acceso a ellos. Ya señalé que estas distinciones sólo están disponibles para un número reducido de teóricos especializados. Aunado a lo anterior, también ya señalé que los teóricos especializados como Billon y Kriegel están al tanto de que la fenomenología no se define por el registro o, en este caso, por la capacidad de asimilar o comprender información, sino por cómo se siente para el sujeto estar en cierto estado. Luego, también es un error concluir la ausencia de fenomenología a partir del hecho de que el sujeto no puede comprender la información de lo que ve o de lo que oye, por ejemplo.

Por último, los negacionistas también parecen asumir que otros reportes donde los sujetos expresan consciencia e inconsciencia refieren a múltiples estados fenoménicamente inconscientes (o que *la mayoría* de los estados son fenoménicamente inconscientes), pero accesibles, pues señalan lo siguiente: "los pacientes explícitamente afirman que sus estados

[25] Véase la sección 2.2 sobre la función de la consciencia de acceso. Su función es poner la información a disposición de otros sistemas.

alienados son inconscientes aunque son conscientes *de* ellos" [Billon y Kriegel, 2015, p. 45]. Para ilustrar su punto, citan el siguiente reporte: "de pronto me pregunto: ¿realmente estoy aquí? ¿Realmente soy yo quien camina? Entonces hago enormes esfuerzos por aplicar mi consciencia a esta inconsciencia… para darme cuenta de que yo estoy haciendo los movimientos andantes. Entonces, en algún punto durante esta especie de crisis, antes de la absoluta certeza (de ser yo mismo, antes de la crisis) soy consciente por un lado de que soy inconsciente por otro lado" [Séglas y Meige, 1895, p. 147].

No obstante, el reporte no parece apuntar a múltiples estados alienados. En cambio, parece referir a que el sujeto es inconsciente (supongamos que fenoménicamente) de que él mismo está realizando los movimientos andantes. Pero, aun si concedemos esto último, no parece referirse ni siquiera a un solo estado en particular y/o no es claro cuál sería el estado que es fenoménicamente inconsciente. Y, de nuevo, lo que la teoría predice es que el sujeto es consciente de varios estados inconscientes o de un estado específico inconsciente.

Ante esta dificultad, la teoría podría suponer que este caso es similar al que vimos en la sub-sección 1.4.2 del primer capítulo, a saber, en el que el sujeto reporta que sus manos y pies se mueven automáticamente (no siente o no le parece que él los mueve o los controla). Vimos que la teoría gradualista supone que, en este caso, el estado relevante o aquél que subyace la sensación de automatización, es una intención para la acción y, he señalado que la teoría negacionista también supone lo anterior.

Ahora, si asume lo anterior en relación con ese reporte, entonces también tendría que suponer que en este caso es una intención para la acción la que no es fenoménicamente consciente y, adicionalmente, que el paciente tiene acceso a dicha intención. Desafortunadamente, este supuesto es poco plausible si consideramos que nada en el reporte sugiere que el paciente tenga una intención que no se sienta suya y que lo incite a actuar de todas formas. En otras palabras, el reporte no sugiere algo como lo siguiente: "la intención no es mía, pero esta me incita o *me informa* cómo moverme". Más aún, podríamos contrastar este tipo de reporte con los delirios de control que expresan los pacientes esquizofrénicos. En estos casos, los pacientes suelen reportar que están sujetos a la voluntad de alguien más e incluso que las voces en su cabeza los obligan a actuar de ciertas formas (véase DSM-V y Billon y Kriegel, 2015). Quizás en estos casos puede parecer que sí hay una intención que no

se siente propia y a la que tiene acceso el sujeto (aunque habría que revisar los detalles), pero nuestro caso de despersonalización no es similar[26].

No hay rastro de ausencia de subjetividad: Respecto al aspecto fenoménico, ya hemos visto que otra predicción de la teoría es que, si un estado no es subjetivo, entonces el sujeto expresará que no se siente como si fuera suyo (por lo pronto dejemos de lado el aspecto del acceso) y, en consecuencia, dicho estado no es fenoménicamente consciente. Sin embargo, algunos reportes parecen apuntar en otra dirección.

Por ejemplo, ya señalé en el primer capítulo (véase 1.4.2) que algunos pacientes reportan emociones aplanadas o disminuidas y los mismos negacionistas reconocen que es común que los pacientes reporten aplanamiento de emociones. Para ilustrar su punto recurren al siguiente reporte: "la parte emocional de mi cerebro está muerta… todas mis emociones están aplanadas" [Shorvon, 1946, p. 784]. Como también mencioné en el primer capítulo, en este caso el sujeto no está sugiriendo que sus emociones no se sienten como si fueran suyas, lo que está expresando es que *sus* emociones están atenuadas o disminuidas.

Ahora bien, si la teoría negacionista postula que el estado no es fenoménicamente consciente debido a que no es subjetivo y, si el sujeto no está diciendo que sus emociones no se sienten suyas, sino aplanadas ("todas *mis* emociones están aplanadas") entonces no tenemos razones para pensar que sus emociones no son fenoménicamente conscientes.

Anteriormente, vimos que una hipótesis que podía añadir la teoría gradualista respecto a las emociones disminuidas era que, si la subjetividad de una emoción disminuía, entonces la emoción disminuía en intensidad. Sin embargo, la teoría negacionista no parece tener los recursos para dar cuenta de una emoción disminuida pues, de inicio, es una alternativa a la gradualista. Además, explícitamente afirma que el estado no es fenoménicamente consciente en absoluto, por lo tanto, excluye los grados. Sin embargo, de

[26] Por otro lado, aún si asumiéramos que el estado fenoménicamente inconsciente es una intención, el reporte del paciente iría en contravía de lo que afirma la teoría. Recordemos que los negacionistas sostienen que los pacientes no tienen que hacer ningún esfuerzo para acceder a sus estados alienados. Sin embargo, el sujeto expresa que tiene que hacer *enormes esfuerzos* para aplicar su consciencia (supuestamente de acceso) a dicha inconsciencia (o a la supuesta intención que es fenoménicamente inconsciente).

incorporar aspectos de la teoría gradualista, heredaría las dificultades que mencioné en el capítulo anterior.

Otros pacientes expresan que se sienten desapegados de ciertos estados. Los siguientes reportes ilustran estos casos: “me siento tan desapegado de mis pensamientos que parecen tener una ‘vida’ propia” [Sierra y Berrios, 2000, p. 163] y “cuando una parte de mi cuerpo me duele, me siento tan desapegado del dolor que se siente como si fuera el dolor de alguien más” [Ibíd.]. Siguiendo la postura negacionista, en estos casos los estados que no serían subjetivos y, en consecuencia, fenoménicamente inconscientes, serían los pensamientos y los dolores. No obstante, el desapego suele definirse como falta de interés o indiferencia hacia algo y es diferente sentir indiferencia hacia algo y sentir que algo no es mío o como si no me perteneciera, que es lo que predice la teoría.

Si es diferente, entonces la teoría tendría que recurrir a una hipótesis adicional para explicar el desapego. Esta hipótesis podría ser la siguiente: si los sujetos sienten indiferencia hacia los estados es porque, aunque no lo expresen explícitamente, estos no se sienten como si fueran de ellos. Es decir, aunque sean aspectos diferentes, están relacionados, de manera que si el estado no se siente propio, entonces produce indiferencia hacia él.

Sin embargo, esto no parece ser el caso. El primer sujeto reporta que se siente desapegado de *sus* pensamientos, mientras que el segundo expresa que una parte de *su* cuerpo le duele, lo que sugiere que siente *su* dolor corporal o que es él mismo quién siente dolor. Además, recordemos que los mismos negacionistas reconocen que los pacientes no tienen fallos de auto-atribución. Es decir, se atribuyen los estados como suyos. Siendo así, si los pacientes sienten que son sus dolores o pensamientos, aunque sientan indiferencia hacia ellos, entonces la hipótesis no resulta plausible y no tenemos razones para pensar que los estados carecen de subjetividad y, por tanto, de fenomenología propia.

Ahora, recordemos que, en la sección 2.2, señalé que los negacionistas sostienen que los pacientes despersonalizados aún creen que los estados alienados son suyos. De ser así, podrían replicar que, aunque los pacientes reconocen a nivel cognitivo (lo saben o lo creen) que son sus pensamientos o dolores, no los sienten suyos a nivel experiencial; por esta razón, dicen que parecen tener una vida propia o que se siente como si fuera el dolor de alguien más (no suyo). Para esto tendríamos que suponer que, “por vida propia”, el sujeto se refiere

inequívocamente a "no son míos o no se sienten míos". Si bien podemos conceder que es posible, también hay diferentes maneras en las que se puede entender la expresión. Por ejemplo, cuando no quieres pensar algo en específico y el pensamiento surge recurrentemente, una expresión adecuada es que parece tener vida propia. No es que no sea mi pensamiento, es mío, pero llega sin que yo quiera.

Sin embargo, para el caso del dolor, el reporte parece apuntar en otra dirección. Ya vimos que el sujeto sugiere que tiene dolor o que él siente su dolor corporal: "cuando una parte de *mi* cuerpo *me* duele", lo cual va en contravía de la sugerencia según la cual el sujeto sólo cree que es su dolor sin sentirlo suyo o sin sentir que él mismo lo tiene o lo experimenta.

Respecto a los dolores, también hay pacientes que expresan que no les parecen genuinos. Por ejemplo, "era doloroso y sentí que mi brazo se retraía, pero no era un dolor genuino, era un dolor que no llegaba al alma… es un dolor, si quieres, pero la superficie de mi piel está tres kilómetros lejos de mi cerebro y no sé si estoy sufriendo" [Janet, 1928, p. 65]. En este caso tampoco encontramos nada que sugiera que el dolor del paciente no se siente suyo o que no siente dolor. Aun así, de manera similar a los casos anteriores, la teoría podría suponer que, aunque el paciente no exprese explícitamente que no es su dolor, este no se siente propio y, al no sentirse de esta manera, no le parece genuino o real.

No obstante, dicho supuesto también es implausible por la siguiente razón: si el estado no es fenoménicamente consciente, entonces no hay nada que se sienta para un sujeto estar en un estado o tener una experiencia, pero, de hecho tenemos razones para pensar que el sujeto *siente* dolor, pues reporta "era doloroso", entonces precisamente *sentía* que era doloroso. También reporta "es un dolor, si quieres", lo cual sugiere que reconoce que siente algún tipo de dolor.

Así, si siente dolor, entonces hay algo que se siente para el sujeto estar en un estado de dolor o tener una experiencia de dolor. Luego, el estado es fenoménicamente consciente y, por lo tanto, es subjetivo.[27] En consecuencia, este supuesto no solo es implausible, sino que va en contra de lo que reporta el paciente.

[27] Cabe señalar que los negacionistas también tratan de interpretar los reportes de los pacientes esquizofrénicos que padecen de pensamientos insertados (el delirio de que tienen pensamientos en ellos o en su cabeza que no son suyos) en términos de acceso sin fenomenología. En este caso, sostienen que hay

La teoría no es específica respecto al acceso: En términos generales, los reportes citados anteriormente apuntan a que no tenemos buenas razones para pensar que los estados de los sujetos son fenoménicamente inconscientes, lo cual bastaría para poner en duda la predicción de la teoría según la cual los pacientes tienen acceso a estados que *carecen de fenomenología propia*.

Sin embargo, aun si lograra mostrar que los estados no son fenoménicamente conscientes, la teoría no especifica cómo o de qué manera el sujeto puede acceder a la información que transmite el estado, pues se enfoca precisamente en tratar de mostrar la ausencia de fenomenología. Por ejemplo, respecto a algunos reportes, lo único que mencionan los negacionistas es que, "algunos (pacientes) sugieren que no hay nada que se sienta para ellos ver (…) o incluso que no hay nada que se sienta percibir" [Billon y Kriegel, 2015, p. 45. Paréntesis mío]. Para ilustrar su punto recurren a los siguientes reportes: "todo en la visión está muerto" [Mayer-Gross, 1935, p. 111] y "puedo ver, oír y oler, pero es como si no viera, oyera u oliera." [Dugas y Moutier, 1911, p. 10].

Si bien no están sugiriendo nada respecto al acceso[28], sabemos ya que la predicción de la teoría es que, aunque dichos estados no son fenoménicos, son accesibles al sujeto. Esto es, el sujeto puede reportar la información que transmiten. No obstante, en primer lugar, en estos casos los sujetos están reportando estados visuales[29], auditivos y olfativos. Pero estos

---

un estado de orden superior que *es fenoménicamente consciente* y que representa como "en él" (en la cabeza del paciente) un estado de orden inferior que *no es fenoménicamente consciente*. Para Billon (2013), por ejemplo, esto ejemplifica una fenomenología "extrínseca" o de orden superior. "La fenomenología de estar presentes (en él) de una manera que es fenoménicamente consciente." [2013, p. 16. Paréntesis mío]. Si tomamos en cuenta lo anterior, podríamos pensar que para dar cuenta de la fenomenología en el caso de los pacientes despersonalizados, los negacionistas podrían afirmar que los estados de orden inferior de estos pacientes no son fenoménicamente conscientes (y esto es lo que importa, pues, presuntamente, el estado al que se accede o se reporta es el que no es fenoménico), pero que hay una fenomenología extrínseca o de orden superior que hace que los estados se sientan en ellos (pero no suyos). No obstante, la opción tampoco parece viable porque no encontramos en los reportes nada parecido. Es decir, a diferencia de los pacientes esquizofrénicos, los pacientes despersonalizados no reportan ni sugieren que tienen estados *en ellos* que no son suyos.

[28] Recordemos, además, que lo único que dicen sobre el acceso es que los sujetos despersonalizados acceden a sus estados alienados *de manera similar* a como nosotros accedemos a nuestros estados fenoménicamente conscientes (véase sección 2.2). Luego, cabría esperar que haya diferencias entre la manera en la que se accede a los estados. Sin embargo, tampoco dicen nada respecto a este punto.

[29] En el caso del paciente que reporta que todo en la visión está muerto, podríamos interpretar que el sujeto ve, pero todo lo que ve está muerto o parece muerto. Y, de hecho, hasta donde sabemos, el sujeto puede ver. Luego, también estaría reportando estados visuales.

reportes son consistentes con la tesis según la cual los sujetos reportan estados perceptivos que se sienten de cierta manera. Es decir, que reportan estados fenoménicos o experiencias perceptivas.

En segundo lugar y, relacionado con el punto anterior, si dichos estados no se sintieran de cierta manera o si no hubiera nada que se sienta para los sujetos ver, oír u oler, cabría esperar que reportaran que no pueden ver, oír u oler y que de alguna manera saben cuál es el contenido de esos estados. Pero esto no es lo que encontramos. De ser así, la teoría tendría que proporcionar una concepción del acceso sin recurrir a experiencias o estados que se sienten de cierta manera. En la siguiente sección voy a abordar detalladamente estos problemas.

Por lo pronto y, de manera general, hemos visto que la teoría predice que si un estado o múltiples estados de los sujetos no son fenoménicos, pero tienen acceso a ellos, entonces estos reportarán que no se sienten como si fueran suyos, pero reportarán la información que transmiten dichos estados o, alternativamente, que si un estado o múltiples estados no son fenoménicos, pero estos son accesibles, entonces reportarán que son conscientes en un sentido, pero no en otro. Esto es, que son conscientes (en el sentido de acceso) de sus estados inconscientes (en el sentido fenoménico).

No obstante, de seguir la sugerencia de tomarlos de manera literal, señalé que la teoría enfrenta dificultades interpretativas debido a que muchos de los reportes no confirman las predicciones de la teoría y/o parecen apuntar en otra dirección. Relacionado con estos problemas generales, señalé, en primer lugar, que aunque algunos sujetos reportan que son conscientes en un sentido, pero no en otro, no es claro que se refieran a los sentidos de consciencia a los que apela la teoría. En segundo lugar, muchos de los reportes no sugieren que los estados de los pacientes no sean subjetivos y, por tanto, que no sean fenoménicamente conscientes. En tercer lugar, la teoría no especifica cómo se puede acceder a la información de un estado sin recurrir a experiencias o estados que son fenoménicamente conscientes.

En vista de estos problemas, argumenté que para que la teoría pudiera solucionar estas dificultades y/o para que los reportes cumplan con las predicciones de la teoría, esta debería ser enriquecida con hipótesis auxiliares. Sin embargo, mostré que estas son problemáticas debido a que resultan poco plausibles o debido a que van en contra de lo que postula la teoría

o en contra de los mismos reportes. En otros casos, no encontramos buenas razones para pensar que los estados que reportan los sujetos carecen realmente de fenomenología. Por último, respecto a la tercera dificultad, señalé que en la siguiente sección voy a desarrollarla de manera detallada. En este caso no se trata de añadir hipótesis auxiliares, sino, precisamente, la dificultad reside en que la teoría no explica cómo se da el acceso en estos casos.

2.4. Dificultades explicativas respecto al acceso

En el apartado anterior vimos que, a nivel interpretativo, no parece haber nada que sugiera que los pacientes despersonalizados tienen acceso a estados fenoménicamente inconscientes, pues algunos de los estados que reportan son consistentes con la tesis según la cual reportan estados perceptivos que se sienten de cierta manera. Además, mencioné que, si no se sintieran de cierta manera, cabría esperar que los sujetos reportaran que no pueden percibir o que son incapaces de hacerlo.

Para ver por qué deberíamos esperar lo anterior, voy a recurrir al caso hipotético que plantea Block (1995/2007) para ilustrar la posibilidad conceptual de acceso sin fenomenología. A partir de este caso, voy a argumentar que la teoría enfrenta dificultades explicativas pues, por una parte, si los negacionistas están tomando las distinciones de acceso y fenomenología de Block, entonces podríamos esperar que los reportes de los pacientes sean similares al caso hipotético de Block. Sin embargo, esto no es lo que encontramos. Por otra parte, al no postular un mecanismo de acceso similar al que plantea Block en su caso hipotético, la teoría no ofrece una concepción del acceso que no explote la fenomenología del estado al que se accede. En la sub-sección 2.4.1, presentaré el caso hipotético de la súper visión ciega de Block y en la sub-sección 2.4.2, abordaré las dificultades mencionadas previamente.

2.4.1. Acceso sin fenomenología: el caso de la súper visión ciega

Como hemos visto hasta ahora, un estado fenoménicamente consciente es un estado tal que hay algo que se siente para un sujeto estar en él. Por otro lado, en la sección 2.2, vimos que Block, quien originalmente planteó estas distinciones entre fenomenología y acceso, sostiene que la consciencia de acceso tiene una función y que esta es “permitir de alguna manera que la información representada en el cerebro se utilice para razonar, informar y guiar

racionalmente la acción" [Block, 1995/2007, p. 160]. Esto es, un estado es A-consciente si la información que porta el estado está disponible para su uso en el razonamiento, el control racional de la acción y el habla.

Si, como sostiene Block, estos dos diferentes conceptos refieren a propiedades independientes (aunque *normalmente* interactúan o se ejemplifican al mismo tiempo), entonces, siguiendo al autor, es posible que pueda instanciarse una sin la otra. Es decir, consciencia fenoménica sin consciencia de acceso[30] y consciencia de acceso sin consciencia fenoménica, que es lo que postula la teoría negacionista.

Un caso de acceso sin fenomenología podría ejemplificarse con los zombies filosóficos. Como vimos en la sección 2.3, un zombie filosófico sería una criatura (hipotética) que no tiene ningún estado fenoménicamente consciente, pero sí acceso a la información que transmiten. Es decir, esta criatura sería capaz de reportar la información de sus estados fenoménicamente inconscientes, así como de realizar una serie de acciones y formarse pensamientos o creencias a partir de la información que transmiten dichos estados (para detalles sobre la posibilidad metafísica de estas criaturas, véase Chalmers, 1996).

Sin embargo, Block descarta este caso hipotético debido a que está al tanto de que la posibilidad metafísica de que existan estas criaturas *sin ningún* estado fenoménicamente consciente es ampliamente controvertida. Así, plantea otro caso hipotético de un zombie parcial que consiste en afirmar que *solo algún* estado es fenoménicamente inconsciente, pero accesible. Siguiendo al autor, este caso sería menos controvertido[31].

Para entender este caso, conviene introducir primero el caso real de una condición conocida como *visión ciega*. Los sujetos que padecen de esta condición tienen un daño en la corteza visual primaria y, como consecuencia, tienen áreas ciegas en su campo visual. Como es de esperarse, si se les presenta un estímulo en esta área ciega, los sujetos afirman que no ven nada y, según Block, si no ven el estímulo, entonces no hay nada que se sienta para ellos

[30] Si bien no me ocupo de la posibilidad de fenomenología sin acceso, puede consultarse Block (2007,2011) y Sebastián (2014) para ver en qué consiste.

[31] Como también vimos en la sección 2.3, los negacionistas suponen que, cuando los pacientes se describen como zombies, se refieren a que son zombies filosóficos. No obstante, al igual que Block, plantean solo zombies parciales. Recordemos que lo que postula la teoría es que los pacientes despersonalizados solo tienen *algunos* estados fenoménicamente inconscientes y, cuando la despersonalización alcanza su clímax, *la mayoría* (pero no todos) los estados de los pacientes dejan de ser fenoménicamente conscientes.

verlo[32]. Sin embargo, si se les obliga a adivinar entre un número limitado de opciones qué estímulo se les presentó, *algunos* sujetos logran informar la identidad del estímulo o algunas características de él. De la misma manera, si se les incita a tomar un objeto que afirman que no ven, suelen acomodar la mano de una manera adecuada para tomarlo[33].

El caso de la visión ciega no logra ilustrar la independencia de la consciencia fenoménica y la consciencia de acceso, pues, aunque los pacientes no tienen un estado visual fenoménicamente consciente, la información de ese estado tampoco está disponible para el control racional de la acción y el habla[34]. Recordemos que en la sección 2.2 señalé que, para Block, solo hay control racional si se actúa o se reporta algo de manera voluntaria o espontánea o si hay una razón o motivo para hacerlo. Ahora, incluso si estos pacientes son capaces de adivinar la identidad del estímulo que no pueden ver y pueden actuar o iniciar una acción cuando se les solicita, ninguna de estas tareas es espontánea o voluntaria. Actúan y reportan solo cuando otra persona los obliga o los incita a hacerlo. Luego, el estado (la información que porta) no cumple con las características para ser A-consciente y, por tanto, la visión ciega no es un caso de acceso sin fenomenología.

No obstante, Block plantea el caso hipotético de la *súper visión ciega*. En este caso, el supuesto paciente con súper visión ciega ha sido entrenado para obligarse a sí mismo a adivinar qué hay en su campo ciego. Más aún, a diferencia de un paciente con visión ciega al que le piden adivinar qué estímulo está presente a partir de un número limitado de opciones, el paciente con súper visión ciega se obliga a adivinar sin que le presenten un conjunto limitado.

El resultado de esta incitación voluntaria es el siguiente: "El paciente con súper visión ciega *espontáneamente* dice 'ahora sé que hay una línea horizontal en mi campo ciego aunque en realidad no la veo'. La información visual de su campo ciego simplemente aparece en sus pensamientos de la misma manera (…) en que algunas personas simplemente saben la hora

[32] Esta interpretación es controvertida. Algunos autores, como Phillips (2018), han argumentado que no tenemos razones para pensar que hay percepción inconsciente (en el sentido fenoménico). Sin embargo, *por mor* del argumento, acepto la interpretación que ofrece Block.
[33] Para más detalles del fenómeno, véase Weiskrantz (1986), Farah (1994) y Block (1995/2007).
[34] Block sostiene que tampoco estaría disponible como premisa para un proceso de razonamiento, a menos que el paciente escuchara y creyera su propia conjetura.

o dónde está el norte sin tener una experiencia perceptiva de ello" [Block, 1995/2007, p. 173. Énfasis mío.].

Así, aunque el estado del paciente con súper visión ciega no es fenoménicamente consciente, es A-consciente, pues la información está disponible para que el sujeto se forme un pensamiento o creencia, a la vez que está disponible para reportarla de manera espontánea o voluntaria. De acuerdo a Block, este sería un caso menos controvertido de acceso sin fenomenología y un caso de zombie parcial.

2.4.2. Dificultades a partir del caso hipotético de Block

Ahora bien, si los negacionistas están tomando las distinciones de acceso y fenomenología de Block, entonces podríamos esperar que los casos de despersonalización sean similares al caso hipotético de la súper visión ciega. Sin embargo, esto no es lo que encontramos.

Consideremos los últimos dos reportes que vimos en la sección 2.3: "todo en la visión está muerto" y "puedo ver, oír y oler, pero es como si no viera, oyera u oliera". Si estos reportes sugirieran acceso sin fenomenología, dado el caso de la súper visión ciega, cabría esperar que los sujetos reportaran que no pueden ver, no pueden oír o no pueden oler, pero que de alguna manera *saben* cuáles son los contenidos de estos estados o saben qué información transmiten. Ahora, no solo no encontramos nada parecido en los reportes de estos pacientes, sino que, como he señalado, al reportar estados visuales, auditivos y olfativos, los reportes son consistentes con la tesis según la cual reportan estados perceptivos que se sienten de cierta manera.

Por otra parte, parece que el caso de la súper visión ciega debe su plausibilidad a un mecanismo alternativo de acceso que consiste en que, *espontáneamente, el sujeto adivina* (o le permite adivinar de manera espontánea) qué hay en su campo ciego. En otras palabras, el paciente con súper visión ciega sabe que hay una línea horizontal en su campo ciego, aunque no la ve, debido a que hay un mecanismo alternativo de acceso que le permite adivinar espontáneamente que hay una línea horizontal.

Sin embargo, relacionado con el punto anterior, los reportes de los pacientes no sugieren que utilicen un mecanismo alternativo de acceso que les permita *adivinar* espontáneamente la información de sus estados perceptivos. Si bien podemos conceder que

el mecanismo alternativo de acceso no tiene que ser exactamente el mismo que en el caso de la súper visión ciega, la teoría tendría que proporcionar un mecanismo alternativo análogo que les permita a los pacientes reportar sus estados que, presuntamente, carecen de fenomenología. Desafortunadamente, no ofrece ninguno y, de no postular dicho mecanismo, no proporciona una concepción del acceso que no explote la fenomenología del estado al que se accede.

Este problema todavía es más evidente si consideramos otro tipo de reportes. Por ejemplo, aquellos que refieren a dolores. En la sección 2.3, vimos que uno de los pacientes reporta lo siguiente: "cuando una parte de mi cuerpo me duele, me siento tan despegado del dolor que se siente como si fuera el dolor de alguien más." Veíamos también que una opción que tenía la teoría negacionista para dar cuenta de este reporte era suponer que el paciente podía reconocer a nivel cognitivo que el dolor es suyo, pero que no lo siente de esta manera a nivel experiencial.

Si como el paciente hipotético con súper visión ciega *sabe* que hay una línea horizontal en su campo ciego debido a que tiene acceso a la información que transmite el estado fenoménicamente inconsciente, entonces, de manera similar, el paciente despersonalizado podría *saber* que hay dolor o saber que es su dolor porque tiene acceso a la información que transmite este estado que no se siente de ninguna manera para él.

Sin embargo, ante este caso, cabría preguntar cuál es la información a la que el sujeto tiene acceso para que le permita saber que hay dolor y/o qué tipo de información transmite este estado[35]. Supongamos que el estado transmite que hay daño en el brazo, pero, ¿cómo me informa el estado que hay daño en el brazo? Una explicación natural diría que *sintiendo* el daño en el brazo o que el estado informa que hay daño en el brazo al *hacerme sentir de alguna manera*. Precisamente, al hacerme sentir dolor.

Relacionado con este punto, hay autores que afirman que un sujeto solo puede saber que tiene dolor al sentirlo. Por ejemplo, Aydede señala lo siguiente: "los dolores son algo

[35] El caso del dolor es complejo, pues es ampliamente aceptado que este estado está compuesto por diferentes características o cualidades (véase, por ejemplo, Aydede, 2019). Siendo así, podría ser difícil establecer una característica en específico a la cual se podría acceder. Para mantener las cosas simples, podemos asumir que una de ellas es un daño corporal o un daño en una localización corporal.

privado de sus propietarios, en el sentido estricto de que nadie más puede acceder epistémicamente al dolor de uno mismo de la misma manera en la que uno tiene acceso a su propio dolor. Es decir, sintiéndolo y llegando a saber que lo está sintiendo a partir de esa base" [Aydede, 2019]. Más aún, añade que, "las personas no solo parecen tener un acceso epistémico especial a sus propios dolores, sino que también parecen tener una *autoridad* epistémica muy especial con respecto a su dolor: parecen ser *incorregibles* o, incluso infalibles en cuanto a sus dolores y sus reportes de dolor (…) si siento dolor, entonces sé que tengo dolor[36]" [Ibíd.]

Ahora bien, la teoría no solo niega la explicación natural del acceso a los dolores, pues, precisamente postula que el sujeto tiene acceso a la información que transmite un dolor (por eso sabe que hay dolor) que no se siente de ninguna manera para el sujeto, sino que tampoco ofrece una explicación alternativa del acceso que podamos evaluar al respecto. Al no ofrecer una explicación alternativa, de nuevo, no proporciona una concepción del acceso que no explote la fenomenología del estado al que se accede. Esto es, sin lugar a duda, un problema serio para el negacionismo.

En resumen, en esta sección argumenté que la teoría enfrenta serias dificultades explicativas, pues si los negacionistas están tomando las distinciones de acceso y fenomenología de Block, deberíamos encontrar que los casos de despersonalización son similares al caso hipotético de la súper visión ciega que plantea Block para ilustrar la posibilidad conceptual de acceso sin fenomenología. Sin embargo, no encontramos nada parecido en los reportes. Más aún, dado que la teoría no ofrece una explicación alternativa del acceso o un mecanismo alternativo de acceso análogo al que plantea Block en su caso hipotético, no proporciona una concepción del acceso que no explote la fenomenología del estado.

[36] Podría objetarse que los sujetos no son infalibles respecto a sus dolores y, en general, respecto a sus experiencias y reportes de ellas. Ante esto podríamos decir que, *en condiciones normales*, hay autoridad epistémica respecto a nuestras propias experiencias. No obstante, también puede haberla en condiciones patológicas. Por ejemplo, hemos visto que en el caso del dolor del paciente despersonalizado, teníamos razones para pensar que era fenoménicamente consciente, lo cual iba en contravía de la sugerencia de que el sujeto solo sabía o creía que era su dolor *sin* sentirlo. Luego, la observación de Aydede puede reforzar el punto de que el sujeto sabía que era su dolor debido a que lo sentía.

Si tomamos en conjunto las dificultades interpretativas y explicativas, estas apuntan, principalmente, a que no tenemos buenas razones para pensar que los pacientes despersonalizados tienen acceso a estados que carecen de fenomenología propia.

En la última sección, voy a mostrar que, además de enfrentarse a las dificultades previamente mencionadas, la teoría también se enfrenta con dificultades empíricas pues, en términos generales, la evidencia empírica disponible tampoco parece converger con lo que plantea la teoría.

2.5. Dificultades con la evidencia empírica

Si la teoría postula ausencia de fenomenología de algunos estados y acceso a la información que transmiten, entonces la evidencia neurológica debería mostrar un buen funcionamiento del acceso, pues recordemos que los defensores de la teoría señalan que los pacientes no tienen que hacer *ningún esfuerzo* para acceder a sus estados fenoménicamente inconscientes, lo que sugiere que, en general, no tienen mayores dificultades para acceder a sus estados. Por otra parte, deberíamos encontrar daños en la fenomenología.

No obstante, la evidencia empírica respecto a la fenomenología es un asunto delicado y complicado debido a que, en la actualidad, no se sabe a ciencia cierta cuáles son los correlatos neuronales que subyacen a la consciencia fenoménica. Si bien este puede ser un problema, voy a señalar que, a partir de la misma evidencia, podemos hacer ciertas asunciones sobre los correlatos que la pueden subyacer.

En la sub-sección 2.5.1, voy a presentar las áreas cerebrales que se han asociado con el acceso, así como el rol de la atención en el acceso. De manera general, la atención parece ser una facultad o capacidad de suma relevancia para que el contenido de un estado pueda accederse. Una vez identificados estos aspectos, voy a presentar pruebas de neuroimagen y pruebas neuropsicológicas que, en conjunto, parecen apuntar a daños en el acceso y/o a que hay dificultades con el acceso. Posteriormente, en la sub-sección 2.5.2, voy a señalar que, pese a que no existe una idea clara respecto a las áreas que subyacen a la fenomenología de los estados, podríamos asumir que ciertas áreas involucradas en las percepciones o en las emociones, por ejemplo, *contribuyen* a la fenomenología de dichos estados. De esta manera, si hay daños en dichas áreas, esto podría mostrar daños en la fenomenología. No obstante, aun asumiendo lo anterior, la evidencia no parece converger con los postulados de la teoría.

2.5.1. Consciencia de acceso y evidencia empírica

Respecto a las áreas cerebrales asociadas con el acceso, Block (2007, 2011) ha planteado que el acceso está asociado a la memoria de trabajo y que los mecanismos subyacentes a esta se asocian con áreas frontales del cerebro, específicamente con la corteza pre-frontal dorsolateral (CPDL).

Por otro lado, en el ámbito neurocientífico, la teoría más elaborada y reconocida de la consciencia de acceso es *la teoría del espacio de trabajo global* (Baars, 1988; Dehaene, 2009; Dehaene et al. 2006) según la cual "un estado es consciente cuando y solo cuando él (o su contenido) está presente en el espacio de trabajo neuronal global, lo que hace que el estado (contenido) sea globalmente *accesible* a múltiples sistemas" [Wu y Morales, 2024].

En términos generales, los defensores de esta teoría sostienen que el espacio de trabajo neuronal global está constituido por conexiones neuronales de largo alcance y que "la conectividad a larga distancia de estas 'neuronas del espacio de trabajo', cuando están activas durante un periodo mínimo, puede hacer que la información esté disponible para una variedad de procesos, incluida la categorización perceptiva, la memorización a largo plazo, la evaluación y la acción intencional.[37]" [Dehaene y Naccache, 2001, p. 1.]

Ahora bien, para los defensores de esta teoría, las conexiones neuronales de largo alcance del espacio de trabajo neuronal deben estar distribuidas en al menos cinco áreas cerebrales. Dentro de ellas se encuentran la CPDL, la corteza cingulada anterior y posterior (CCA y CCP), la corteza parietal inferior (CPI), la pre-motora y la temporal superior. No obstante, ponen especial énfasis en la CPDL, la CCA y la CPI. "Las neuronas del espacio de trabajo pueden ser particularmente densas en las cortezas pre-frontales y cingulada anterior, lo que confiere a estas áreas un papel dominante" [Dehaene y Naccache, 2001, p. 14]. Y, "las neuronas piramidales de las capas 2 y 3 (…) aunque se encuentran en toda la corteza, son

[37] Para Dehaene y Naccache, la acción intencional puede ser tanto motora como verbal. Por otro lado, aunque no mencionan nada en específico sobre el razonamiento, postulan que la consciencia es necesaria para combinar diferentes operaciones mentales y usarlas para evaluar y planear una estrategia. La evaluación y la planeación pueden ser procesos involucrados o relacionados con el razonamiento. Por ejemplo, puedo evaluar ciertas creencias que tengo y, a partir de ellas, planear una estrategia argumentativa. Siendo así, al adoptar esta teoría, se dan las características que requiere Block para que un estado sea A-consciente. Recordemos que, para el autor, la información del estado debe estar disponible para su uso en el razonamiento, el control racional de la acción y el habla.

particularmente gruesas en las estructuras corticales pre-frontal dorsolateral y parietal inferior. Entonces, una predicción simple, es que la actividad de estas capas podría estar estrechamente correlacionada con la consciencia"[38] [Ibíd. p. 26].

Así, si bien le atribuyen un papel fundamental a la CPDL para la implementación del espacio de trabajo neuronal global y, en consecuencia, para la implementación del acceso, cuando menos la CCA y la CPI son igualmente relevantes para su implementación.

Por otra parte, a partir de pruebas neuropsicológicas, los defensores de esta teoría han argumentado que *la atención* juega un papel fundamental para determinar si el contenido de un estado es accesible. Por ejemplo, Dehaene y Naccache (2001) sostienen que, para la percepción de un estímulo, no es suficiente que éste esté presente de manera clara y por un periodo considerable de tiempo para que se vuelva consciente, sino que, además, necesita ser atendido por parte del sujeto.

Uno de los fenómenos para ilustrar su punto es aquel que se conoce como *ceguera por falta de atención*. En la ceguera por falta de atención los sujetos no son capaces de percibir un objeto o evento inesperado (claramente visible) debido a que su atención está enfocada en otro estímulo o porque están realizando una tarea que los distrae. Para mostrar este fenómeno, mencionan una prueba neuropsicológica realizada por Mack y Rock (1998) en la que les solicitaron a los sujetos de prueba realizar una tarea de discriminación visual en un área específica de su campo visual. En esa misma prueba, introdujeron un estímulo diferente en otra área de su campo visual. Más aún, sostienen que "este estímulo claramente tenía suficiente contraste y duración (normalmente 200 ms) para ser perceptible de forma aislada" [Dehaene y Naccache, 2001, p. 8]. No obstante, el estímulo pasó desapercibido o no se percibió conscientemente debido a la tarea distractora (discriminación en otra área visual). En consecuencia, la mayoría de los sujetos reportaron que no vieron el estímulo e incluso siguieron negando su presencia cuando se les preguntó si estaban seguros si no había aparecido.

[38] Por motivos de extensión, es imposible describir a detalle los postulados y características de esta teoría. Para más detalles acerca de ella, puede consultar Dehaene y Naccache (2001) y Dehaene, Kerszberg y Changeux (1998).

De esta manera, los autores concluyen lo siguiente: “la orientación de la atención desempeña un papel crucial para determinar si un objeto determinado accede a la consciencia (…) los objetos que no se encuentran en una región atendida del campo visual no pueden ser reportados.” [Ibíd.] Más aún, postulan que el área parietal y la cingulada “contribuyen a la regulación atencional y al cambio de foco de interés (de un estímulo)” [Ibidem, p. 28. Paréntesis mío].

Dado que su teoría es considerada como una teoría de la consciencia de acceso, podemos interpretar su afirmación de manera tal que los objetos que no son atendidos no son accesiblemente conscientes.

De manera similar, Block (1995/2007) ha señalado que aquello a lo que prestamos atención es aquello a lo que podemos acceder o que el contenido de un estado se vuelve accesiblemente consciente con el foco de la atención. Así, si no hay atención a un estímulo visual, por ejemplo, este no entra a (o no está puesto a disposición de) los mecanismos responsables del razonamiento, la acción y el habla.

Siguiendo las consideraciones de Block y de Dehaene y Naccache, Wu (2014) afirma que “el caso canónico de acceso (…) es el acceso a X por parte de los sistemas de memoria de trabajo, donde este acceso está mediado por la atención” [2014, p. 176]. En términos generales, la idea es que para que haya acceso a cierta información X, dicha información debe estar disponible para la memoria de trabajo y, el papel de la atención, es movilizar o entregar dicha información a este sistema. Una vez que la información se encuentra en la memoria de trabajo, esta puede ser accesible (o estar disponible) para guiar el reporte o una acción.

Así, la atención es el mediador entre la información percibida[39] y la memoria de trabajo. De esta manera, si el sujeto no atiende al estímulo X, la información de X no se moviliza a esta memoria y, por tanto, no está disponible para los mecanismos encargados del habla, de la acción y el razonamiento.

[39] La diferencia entre la propuesta de los defensores de la teoría del espacio de trabajo neuronal global y la de Block es que, para los primeros, la percepción es consciente (fenoménicamente), mientras que para Block puede o no puede serlo.

Una vez identificado el papel de la atención en el acceso, así como las áreas cerebrales que lo subyacen, podemos analizar, a partir de pruebas de neuroimagen y neuropsicológicas, si los pacientes despersonalizados no tienen problemas o dificultades con el acceso tal y como postula la teoría negacionista.

En la sección 1.6 del primer capítulo, ya vimos que los estudios de neuroimagen apuntan a daños en la amígdala, hipotálamo e ínsula anterior. Ninguna de estas áreas está asociada con el acceso.

Por otro lado, Sierra y Berrios (1998) han señalado que no han encontrado "informes de casos de despersonalización en el contexto de lesiones pre-frontales" [1998, p. 903]. De manera similar, en un estudio realizado mediante una tomografía por emisión de positrones y mediante una resonancia magnética funcional, Simeon et al. (2000) han señalado que en los sujetos despersonalizados no se encontraron diferencias significativas en la actividad de la corteza pre-frontal, en la corteza cingulada, ni en la corteza pre-central en comparación con los sujetos de control.

Sin embargo, en el mismo estudio, los autores encontraron que, a diferencia de los sujetos de control, los pacientes despersonalizados mostraron más actividad en el área parietal inferior o CPI. Si bien es un área asociada con el acceso, se ha postulado que está involucrada en la *sensación de agencia* de los movimientos. Esto es, la sensación de que uno mismo está realizando ciertas acciones o ciertos movimientos corporales. Por ejemplo, Sierra y David (2011) han señalado que diversos estudios de neuroimagen (Farrer et al. 2003; Farrer et al. 2004 y Farrer y Frith, 2002) han apuntado a que los pacientes que experimentan una falta de sentido de agencia tienen "una activación anormalmente aumentada en el giro angular del lóbulo parietal derecho (…) de modo que cuanto menor sea el sentido de agencia, mayor será la actividad en el lóbulo parietal inferior derecho" [Sierra y David, 2011, p. 104].

Ahora, la sensación de agencia típicamente se considera como un aspecto fenoménico o parte de la fenomenología de algunos estados. Precisamente, los referentes a los movimientos corporales o la sensación de control de ellos. Sin embargo, de tomar en cuenta estos estudios, podríamos pensar que solo mostrarían daños o problemas con la fenomenología. Pero también hemos visto que Dehaene y Naccache postulan que la corteza parietal (al igual que la cingulada) contribuye a la regulación atencional. De ser así, la

actividad anormal de la CPI encontrada en los pacientes despersonalizados, también podría contribuir a problemas atencionales y, en consecuencia, a que cierta información no sea accesiblemente consciente.

Esta hipótesis podría ser probable en tanto se ha señalado que "los pacientes con despersonalización a menudo se quejan de dificultades para enfocar y mantener su atención en estímulos particulares" [Sierra y Berrios, 1998, p. 903] y en tanto se han realizado pruebas neuropsicológicas, como la *prueba de rendimiento continuo de Vigil*, para mostrar déficits de atención en estos pacientes.

En términos generales, en esta prueba se le solicita al sujeto responder cuando aparece un estímulo específico en una pantalla. Por ejemplo, la letra "A". Posteriormente, se le solicita responder solo cuando la letra "A" aparece delante de otra letra o conjunto de letras. Mientras que las otras letras pueden distraer al sujeto de la presencia de la primera, también se puede añadir ruido visual. Por ejemplo, otras letras en pixeles pequeños en la pantalla.

A partir de esta prueba, Guralnik et al. (2000) han mostrado que los pacientes tienen problemas para mantener la atención en una letra específica solo cuando se añade ruido visual. Cuando esto sucede, los autores señalan que los pacientes tienen más errores de omisión que errores o respuestas de comisión. Es decir, al presentar el distractor en la prueba, los pacientes tienden más a dejar de indicar la presencia del estímulo en lugar de identificarlo erróneamente. Precisamente los errores de omisión muestran que, en presencia del distractor, los pacientes son incapaces de fijar la atención en el estímulo relevante y, en consecuencia, no lo indican o no lo reportan.

Tomando en cuenta este estudio y el papel de la atención en el acceso, podríamos decir que el estímulo que no es atendido no se moviliza o no entra a la memoria de trabajo y, por tanto, no está disponible para los mecanismos responsables del habla o de la acción (tanto pueden dejar de reportar la presencia del estímulo, como dejar de presionar un botón cuando este aparece). En otras palabras, la falta de atención interfiere con el acceso al estímulo relevante.

Por supuesto que la atención no es una cuestión de todo o nada. Si bien pueden ser incapaces de fijar la atención en cierto estímulo, también la atención puede simplemente

disminuirse o reducirse. A partir del mismo estudio, Guralnik y sus colegas plantean que la presencia del distractor puede disminuir la sensibilidad perceptiva del estímulo, lo que a su vez provoca que sean incapaces de extraer rápidamente características de él o que les cueste trabajo procesar o asimilar las características relevantes[40]. Así, de manera similar, podríamos plantear que la atención del estímulo se ve reducida por la presencia del distractor y que, en conjunto con la disminución de la sensibilidad perceptiva, esto provoca que les cueste trabajo extraer o recuperar algunas características del estímulo, pero no otras. Siendo así, la atención reducida o disminuida provocaría que haya acceso solo a alguna parte de la información, mientras que otras características de ella permanecerían inaccesibles[41].

En suma, si bien no se han encontrado daños o anomalías en la CPDL o en la CCA, se ha encontrado actividad anómala en la CPI. Si, como plantean Dehaene y Naccache, esta área regula la atención, una hipótesis podría ser que la actividad anómala de esta área contribuye a problemas atencionales.

Relacionado con este punto, las pruebas neuropsicológicas muestran déficits de atención. Si la capacidad atencional se ve comprometida, entonces el estímulo no es accesible a la consciencia o algunas características del estímulo son accesibles y otras no. Luego, los problemas de atención interfieren de una manera u otra con el acceso y, si esto es el caso, es posible que al menos *algunos* pacientes tengan dificultad para acceder a la información de sus estados (de nuevo, por problemas atencionales) o que tengan que hacer un esfuerzo para acceder a ella, así como sugiere uno de los pacientes que cité anteriormente: “hago enormes esfuerzos por aplicar mi consciencia a esta inconsciencia”.

[40] Adicionalmente, los autores también han demostrado que los pacientes tienen problemas con la memoria a corto plazo. Por ejemplo, se les dificulta memorizar figuras geométricas y oraciones largas (o que contienen más carga de información). Los autores postulan que los problemas de memoria a corto plazo derivan de problemas perceptivos (disminución de sensibilidad perceptiva) y atencionales.

[41] Podría objetarse que, en este caso, no hay problemas de acceso porque la información que sí es atendida entra a la memoria de trabajo y, entonces ésta sí está disponible para los mecanismos responsables de la acción y el habla. No obstante, parece razonable pensar que si solo hay acceso a una parte de la información y no a otra, entonces el acceso no es del todo bueno. Por otra parte, si los sujetos reportan características del estímulo que no son relevantes (estas últimas son las que les solicitan reportar) debido a que no prestan atención a las que sí lo son, no parece razonable decir que gozan de un buen acceso porque de cualquier manera están reportando ciertas características, aunque no sean aquellas que se les solicitó reportar.

2.5.2. Consciencia fenoménica y evidencia empírica

Como he señalado anteriormente, la cuestión de la evidencia empírica respecto a la fenomenología es un asunto complicado debido a que no sabemos a ciencia cierta cuáles son las áreas cerebrales que subyacen a la fenomenología o a la consciencia fenoménica. Pero, como también he mencionado, esto no significa que no podamos hacer ciertas asunciones sobre la fenomenología de ciertos estados. Después de todo, acabamos de ver que la sensación de agencia es un aspecto fenoménico de los movimientos o el control corporal y que se ha postulado que la actividad de la CPI está relacionada con dicha sensación. Por otra parte, en la sección 1.6 del primer capítulo, vimos que el sistema límbico y áreas relacionadas están involucradas en la generación de emociones y/o experiencias emocionales.

Siendo así, podría pensarse que la actividad de la CPI *contribuye* a la fenomenología de la agencia y que la actividad del sistema límbico (donde se encuentran la amígdala y el hipotálamo) y áreas íntimamente conectadas con dicho sistema (como la ínsula anterior) *contribuyen* a la fenomenología de las emociones. Es decir, que cuando menos son áreas de suma importancia para que la agencia de los movimientos y las emociones sean estados fenoménicamente conscientes.

Lo mismo podría suceder con otros estados individuales. Por ejemplo, la corteza visual podría contribuir a la fenomenología visual o la corteza temporal podría contribuir a la fenomenología auditiva.

Ahora bien, si la teoría negacionista postula zombies parciales en tanto algunos estados individuales son fenoménicamente inconscientes, entonces podríamos pensar que si el estado que no es fenoménicamente consciente es una emoción, por ejemplo, entonces deberíamos encontrar daños o actividad anómala en las áreas que subyacen estos estados, de nuevo, bajo el supuesto de que estas contribuyen a la fenomenología.

Revisar cada estado y las áreas subyacentes es una tarea ardua que desbordaría este trabajo. No obstante, podemos revisar algunos estados a partir de los estudios que tenemos disponibles.

Comencemos por el sentido de agencia. De nuevo, si como se ha planteado, la ausencia de sentido de agencia se debe a una actividad anormalmente aumentada en el giro

angular del lóbulo parietal, entonces esto mostraría que no hay fenomenología de la agencia o que no hay nada que se siente para el sujeto tener el control de sus movimientos o sentir que él mismo los está realizando.

No obstante, hemos visto que los negacionistas no postulan ausencia de sentido de agencia para dar cuenta de los reportes que refieren a movimientos automáticos. Lo que postulan es que el estado que no es fenoménicamente consciente es *una intención*. Luego, parece que la evidencia no es consistente con lo que postula la teoría. Para que lo sea, la teoría podría asumir que las intenciones para la acción son las que dan lugar al llamado sentido de agencia. De manera que si no hay una intención que se sienta propia y, por tanto, fenoménicamente inconsciente, entonces el sujeto no siente que es el agente de sus movimientos.

Si bien puede ser una opción posible, no está libre de problemas. En primer lugar, ya vimos que, a nivel interpretativo, nada en los reportes de los pacientes sugiere que tengan una intención que no se siente propia (y, adicionalmente, que les informe cómo moverse). En segundo lugar, podríamos cuestionar si la intención es el estado que está relacionado con la actividad de la CPI y si la actividad anormal de esta área deriva en lo que se conoce como falta de sentido de agencia. Es decir, que de alguna manera la intención para la acción no es más que el sentido de agencia mismo. O si la intención y el sentido de agencia son diferentes. De manera que la intención es un estado diferente con áreas subyacentes distintas a la CPI y, quizás al haber actividad anormal en estas áreas, de alguna manera provoca la actividad anormal en la CPI y esto afecta también al sentido de agencia. Hasta que no tengamos una hipótesis clara o evidencia a favor de una de las dos opciones, en principio, la evidencia no converge con lo que postula la teoría.

Pasemos ahora a las emociones y a las áreas asociadas con ellas. En la sección 1.6 del primer capítulo, vimos que en un estudio de resonancia magnética se encontró que en los pacientes despersonalizados no hubo actividad en la ínsula anterior cuando vieron una serie de imágenes aversivas y que, además, reportaron que *no experimentaron ninguna* respuesta emocional al verlas. Lo cual sugería que la ínsula anterior es un área crucial para la generación y experiencia de las emociones.

Si esto es el caso, entonces parece que tenemos dos opciones: 1) hay una emoción o estado emocional que no es fenoménicamente consciente o 2) no surge o no se genera una emoción o un estado emocional. Parece ser que la evidencia apunta a la segunda opción y, de ser así, no es consistente con la predicción de la teoría de que *hay un estado emocional*, pero que este es fenoménicamente inconsciente en tanto no se siente suyo. Más aún, podemos ver que los sujetos reportaron precisamente que no experimentaron ninguna emoción. No que de alguna manera tienen emociones que no se sienten propias.

Por otro lado, vimos que en otro estudio los pacientes solo exhibieron reducción de actividad en la amígdala e hipotálamo en respuesta a caras tristes y felices. Luego, parecería razonable pensar que las emociones o estados afectivos relacionados con la reducción de estas áreas, también están reducidos. Lo cual podría mostrar fenomenología afectiva reducida, no ausencia de fenomenología así como lo plantea la teoría negacionista.

Por último, vimos que los negacionistas suponen que los reportes de algunos pacientes sugieren que no hay nada que se siente para ellos percibir. Si esto es el caso y, si las áreas cerebrales que subyacen la percepción en diferentes modalidades (auditiva, gustativa, olfativa, etc.) contribuyen a la fenomenología perceptiva, deberíamos encontrar daños en dichas áreas. Por ejemplo, consideremos de nuevo a los pacientes con visión ciega y el caso hipotético del paciente con súper visión ciega. En ambos casos, reportan que no pueden ver y en ambos casos se debe a que hay un daño en la corteza visual. Más aún, hemos visto que, para Block, en ambos casos no hay fenomenología visual.

No obstante, en los pacientes despersonalizados no encontramos nada parecido. La evidencia no muestra daños en la corteza visual o en la corteza temporal. En general, parece que los pacientes no tienen problemas perceptivos[42]. Ellos mismos señalan que pueden ver, oír u oler. Siendo así, la evidencia respecto a la percepción no es compatible con la sugerencia de ausencia de fenomenología perceptiva.

[42] Si bien he mencionado que Guralnik et al. (2000) han señalado que los problemas de memoria a corto plazo pueden derivar de *problemas perceptivos* (y atencionales), dicho problema consiste en la disminución de sensibilidad perceptiva del estímulo. Siendo así, podría pensarse que hay disminución de fenomenología visual y que, además, depende del aspecto cualitativo, no del subjetivo. Sin embargo, ya sabemos que la teoría negacionista plantea que la ausencia de fenomenología depende de la ausencia de la subjetividad.

En resumen, la evidencia no converge con la teoría. En primer lugar, si algunos pacientes exhiben déficits atencionales, dichos déficits pueden impedir o dificultar el acceso a la información de los estados. Lo cual sugiere, a diferencia de lo que postula la teoría, que algunos pacientes tienen dificultad para acceder a sus estados. En segundo lugar, al menos algunos estados que revisamos no parecen ser fenoménicamente inconscientes (percepciones), otros podrían implicar disminución de fenomenología afectiva (emociones) y, en otros casos, no parece generarse ningún estado afectivo (si no se genera dicho estado, no puede existir un estado que sea fenoménicamente consciente o inconsciente). Respecto al sentido de agencia, si bien podríamos conceder que una intención fenoménicamente inconsciente podría dar lugar a una falta de sentido de agencia, vimos que la opción no está libre de problemas.

2.6. Conclusión

La teoría negacionista concede que los estados de los pacientes despersonalizados no son subjetivos, pero, de acuerdo al principio subjetivo, si no lo son, entonces tampoco son fenoménicamente conscientes. Sin embargo, la teoría postula que, al ser capaces de notar y reportar dichos estados, los pacientes tienen acceso a ellos. En otras palabras, el problema con los pacientes despersonalizados es que tienen consciencia de acceso a algunos estados que carecen de fenomenología propia.

De esta manera, la teoría predice que, si los estados de los pacientes no son subjetivos y, por tanto, fenoménicamente inconscientes, entonces estos reportarán que no se sienten suyos o como si no les pertenecieran, pero informarán el contenido de dichos estados o, alternativamente, reportarán que son conscientes en un sentido (de acceso), pero no en otro (en el sentido fenoménico). No obstante, he argumentado que, de manera análoga a la teoría gradualista, esta enfrenta una serie de dificultades en diferentes niveles

A nivel interpretativo y siguiendo la sugerencia de tomar los reportes de manera literal, vimos que la mayoría de ellos no cumplían las predicciones de la teoría y, que a partir de este problema general, surgían diferentes dificultades. La primera es que, si bien algunos pacientes expresan explícitamente que son conscientes en un sentido, pero no en otro, no es claro que se refieran inequívocamente a los sentidos de consciencia que plantea la teoría. La segunda es que en muchos reportes no encontramos rastro de ausencia de subjetividad y, si

no hay ausencia de subjetividad, no tenemos buenas razones para pensar que los estados que reportan son fenoménicamente inconscientes. Por último, la teoría no ofrece una explicación del acceso que no dependa de estados fenoménicamente conscientes.

Respecto a las primeras dos dificultades, argumenté que, para resolver dichos problemas, la teoría negacionista debía ser enriquecida con hipótesis auxiliares. Sin embargo, vimos que estas son problemáticas debido a que son poco plausibles o debido a que van en contra de lo que postula la teoría o de los mismos reportes. Si bien la tercera dificultad también se encuentra a nivel interpretativo, señalé que, en este caso, no se trataba de añadir hipótesis auxiliares, sino en que, precisamente, la teoría tendría que ofrecer una concepción del acceso que no explotara la fenomenología del estado. Esto dio lugar a una segunda dificultad.

A nivel explicativo, argumenté que, si los negacionistas están tomando las clásicas distinciones de acceso y fenomenología de Block, entonces deberíamos encontrar que los casos de despersonalización son similares al caso hipotético de la súper visión ciega. Sin embargo, esto no es el caso. Por otro lado, la teoría no ofrece un mecanismo análogo de acceso al que plantea Block en su caso hipotético y, de ser así, no ofrece una concepción del acceso que no explote la fenomenología del estado al que se accede.

Por último, también encontramos dificultades con la evidencia empírica. En este caso señalé que la evidencia no confirma lo que plantea la teoría. Por un lado, la teoría afirma que los pacientes no tienen dificultades con el acceso. Sin embargo, encontramos que las pruebas neurológicas sugieren déficits atencionales que interfieren o dificultan el acceso. Respecto a la fenomenología, vimos que podíamos asumir que las áreas cerebrales involucradas en algunos estados perceptivos, emocionales o aquellas que dan lugar al sentido de agencia, podían contribuir a la fenomenología de dichos estados o sensaciones. No obstante, vimos que, aun asumiendo lo anterior, la evidencia no es consistente con los postulados de la teoría.

Al igual que con la teoría gradualista, mi objetivo no ha sido refutar esta teoría. No niego que, por ejemplo, sea posible que los negacionistas encuentren una manera de mostrar que los estados de los sujetos son fenoménicamente inconscientes. Sin embargo, esto requeriría modificar considerablemente la teoría o ir mucho más allá de lo que originalmente plantea. Asimismo, es posible que puedan encontrar una manera de explicar el acceso a los

estados de los pacientes sin explotar su propia fenomenología. Sin embargo, no parece ser una tarea sencilla.

En esta primera parte del argumento, mostré que cada una de las teorías estructurales enfrenta diferentes tipos de dificultades para dar cuenta de la anomalía de las experiencias o estados que reportan los pacientes despersonalizados. La teoría gradualista enfrenta dificultades interpretativas, teóricas y dificultades con la evidencia empírica. De manera similar, la teoría negacionista enfrenta dificultades interpretativas, explicativas y con la evidencia empírica disponible.

Como he señalado, es posible que ninguna de estas dificultades refute por sí sola cada teoría. Sin embargo, he mostrado que, tomadas en conjunto, son lo suficientemente serias como para buscar una teoría alternativa que nos permita dar cuenta de la anomalía de los estados de estos pacientes.

Lo anterior da paso a la segunda parte del argumento a la mejor explicación. En el último capítulo, voy a esbozar una teoría afectivista que, de manera muy general, plantea que la anomalía de los estados de los pacientes despersonalizados es el resultado de déficits afectivos que pueden manifestarse de diferentes maneras. Esta interpretación afectivista nos permitirá resolver o evitar las dificultades identificadas en el primer paso del argumento.

Capítulo 3

La teoría afectivista

3.1. Introducción

En los capítulos anteriores, vimos que un gran número de reportes sugieren que los pacientes despersonalizados tienen problemas con sus estados afectivos. Por ejemplo, hemos visto que ya no pueden sentir emociones en respuesta a sus familiares o que no pueden experimentar estados de ánimo en general. Asimismo, hemos visto que reportan emociones disminuidas o aplanadas.

Por otro lado, en la sección 1.6 del primer capítulo, vimos que diversos estudios neurológicos han mostrado que, en estos pacientes, hay daños o actividad anómala en áreas cerebrales que se han considerado de suma importancia en la generación de emociones y/o respuestas emocionales, como la amígdala o el hipotálamo.

Si tomamos en cuenta lo anterior, entonces parece razonable pensar que, al menos algunas de las experiencias anómalas de estos pacientes, son causadas por déficits afectivos. Esta idea no es nueva, a partir de los mismos reportes y de los hallazgos neurológicos mencionados previamente, se han ofrecido diferentes teorías afectivistas que pretenden dar cuenta de la anomalía de las experiencias en términos afectivos.

En este capítulo, voy a respaldar esta tesis. Podría pensarse que, al adoptar esta tesis, solo se podrían interpretar algunas experiencias anómalas, precisamente aquellas que refieren a emociones o estados de ánimo. No obstante, al esbozar una teoría afectivista propia, voy a señalar que la mía difiere en un aspecto central. *Grosso modo*, esta teoría plantea que las experiencias anómalas de los pacientes despersonalizados derivan de déficits afectivos que pueden manifestarse de dos maneras: 1) como ausencia o disminución de emociones (o estados afectivos en general) o 2) como ausencia o disminución de dimensiones afectivas específicas, como la valencia y la activación/excitación, que generalmente acompañan a muchos de nuestros estados (no emocionales). La manifestación de los déficits afectivos en estas dos versiones nos permitirá ofrecer interpretaciones alternativas de un buen número de reportes.

Antes de desarrollar esta teoría y de ofrecer una interpretación de los reportes a partir de ella, en la sección 3.2, voy a exponer en qué consisten otras teorías afectivistas. Mi objetivo no es exponerlas a detalle, sino señalar algunos puntos relevantes para ver cómo difieren de mi propia teoría.

En la sección 3.3, voy a desarrollar la teoría afectivista que propongo y voy a señalar que las dimensiones afectivas que plantea, a saber, *valencia* y *activación/excitación*, si bien son aspectos constitutivos de los estados afectivos (emociones y estados de ánimo), no son exclusivas de ellos, sino que, generalmente, otro tipo de estados también incorporan dichas dimensiones.

A partir de esta teoría afectivista, en la sección 3.4, voy a ofrecer una interpretación alternativa de los reportes que consideran las teorías estructurales y voy a mostrar que esta es una mejor interpretación, en tanto no solo evita o resuelve muchos de los problemas interpretativos que encontramos en ambas teorías estructurales, sino que interpreta otros aspectos de los reportes que dichas teorías pasan por alto.

En la sección 3.5, voy a señalar el rol funcional de las áreas cerebrales en las que se han encontrado daños en los estudios realizados a los pacientes despersonalizados y, a partir de esto, voy a mostrar que, a diferencia de las teorías estructurales, la evidencia empírica converge con lo que plantea la teoría afectivista. Esto último respalda que la teoría afectivista puede ofrecer una mejor interpretación de la anomalía de las experiencias que reportan los pacientes con despersonalización. Por último, en la sección 3.6, voy a presentar las conclusiones del capítulo.

3.2. Teorías afectivistas: antecedentes

Como he señalado, la idea de que las experiencias anómalas de los pacientes despersonalizados derivan de déficits afectivos no es nueva. Autores como Gerrans (2015), Sierra y David (2011) y Sierra et al. (2002), han argumentado que estos pacientes exhiben ausencia de *respuestas afectivas* ante personas y situaciones en el mundo que normalmente las causaban. Por ejemplo, Gerrans considera que los pacientes con Cotard padecen de experiencias de despersonalización y/o que los delirios de inexistencia (referentes a ellos

mismos como a objetos del mundo) que expresan estos pacientes son una consecuencia de una experiencia de despersonalización[43]. Así, sostiene lo siguiente:

> el delirio es una respuesta a nivel personal ante una pérdida intratable e impenetrable de respuesta afectiva al mundo. Por supuesto, decir que una experiencia es de despersonalización no es una explicación, sino una caracterización intuitiva: el concepto expresa la fenomenología de sentirse desconectado del mundo, incluido el propio cuerpo (…) Es plausible que tales sensaciones tengan su origen en lo que podríamos llamar desrealización afectiva. El fracaso de los acontecimientos emocionalmente relevantes para desencadenar respuestas afectivas en el paciente, hace que el mundo se sienta irreal y extraño. [Gerrans, 2015, p. 19]

Por otro lado y en una línea similar, Sierra y Berrios (1998) y Sierra et al. (2002) han señalado que la sensación de irrealidad del mundo que suelen reportar los pacientes se explica porque la percepción del mundo (o de los objetos del mundo) no se "colorea" emocionalmente y/o que la percepción no está dotada de "sentimientos emocionales". Esta ausencia de respuesta afectiva la podemos entender como falta de respuesta emocional o, como sugieren Sierra y sus colegas, como ausencia de sentimientos emocionales ante situaciones u objetos del mundo. La idea de falta de respuesta afectiva ante personas es mucho más evidente en tanto es común que experimentemos afecto o emociones específicas hacia ciertas personas, como amigos, parejas, padres y hermanos. En estos casos, la percepción de estas personas tampoco se colorea emocionalmente o no está dotada de sentimientos emocionales.

Ahora bien, la teoría que propongo también postula que, en algunos pacientes, hay una falta de respuesta afectiva hacia personas y eventos u objetos del mundo que normalmente la causaban. Sin embargo, no se restringe a falta de respuestas afectivas o, como señalan Sierra y colegas, a que la percepción no está dotada de sentimientos emocionales. Si la restringimos a estos aspectos, en principio, parecería razonable pensar que sólo podría dar cuenta de experiencias emocionales o estados puramente afectivos.

[43] Si bien distingue ambas condiciones, de manera similar a Billon (2014), afirma que una lleva a la otra. Si esto es el caso o no, no es relevante para este trabajo. Lo importante es que las experiencias de despersonalización las explica en términos de ausencia de respuestas afectivas. Luego, involucran déficits afectivos.

No obstante, como veremos en la siguiente sección, se ha postulado que otro tipo de estados, como pensamientos, sensaciones corporales y percepciones en general, si bien pueden no provocar respuestas afectivas o sentimientos emocionales, sí incorporan dimensiones afectivas específicas como la valencia y la activación/excitación. Siguiendo esta línea de pensamiento, en la siguiente sección, voy a señalar que la teoría que propongo plantea que, si dichos estados incorporan estas dimensiones afectivas, la ausencia o disminución de estas dimensiones pueden provocar que este tipo de estados sean anómalos y, evidentemente, esto nos permitirá dar cuenta de otros reportes que no refieren exclusivamente a emociones o estados de ánimo[44].

3.3. La teoría afectivista y sus dimensiones afectivas: valencia y activación/excitación

Como señalé en la sección 3.1, la teoría afectivista que propongo plantea que las experiencias anómalas de los pacientes despersonalizados derivan de déficits afectivos. Dichos déficits pueden manifestarse de diferentes maneras: 1) como ausencia o disminución de emociones o estados de ánimo o 2) como ausencia o disminución de dimensiones afectivas específicas que, normalmente, acompañan a otro tipo de estados mentales.

Al manifestarse como ausencia de emociones, esta teoría es similar a la de Gerrans y Sierra y colegas, en tanto la ausencia de emociones puede entenderse como ausencia o falta de respuesta afectiva (o emocional) ante objetos, situaciones del mundo y personas[45].

Sin embargo, en la sección 3.2, vimos brevemente que otro tipo de estados pueden incorporar dimensiones afectivas, como la valencia y la activación/excitación. Luego, si hay ausencia o disminución de estas dimensiones afectivas en estos estados, estos pueden ser anómalos.

La cuestión es en qué consisten estas dimensiones afectivas. En esta sección, voy a exponer primero cómo se han entendido en relación con las emociones. Posteriormente, voy

[44] De manera muy general, esta teoría afectivista difiere de las anteriores en que no se restringe a falta de respuestas afectivas, sino que también especifica las dimensiones afectivas mencionadas anteriormente que pueden estar presentes en otro tipo de estados. Sin embargo, también iré mencionando algunas diferencias a partir de la interpretación de algunos reportes.

[45] Aunque, si también conceden que hay emociones aplanadas o disminuidas en los pacientes despersonalizados, entonces la respuesta afectiva o la sensación emocional puede estar disminuida.

a señalar que, pese a que las nociones "valencia" y "activación/excitación" son ambiguas debido a que parecen referir a diferentes aspectos de las emociones, la teoría afectivista adopta una concepción relativamente neutral respecto a ambas. Por último, voy a señalar cómo la valencia y la activación están incorporadas en otros tipos de estados. Una vez que se resuelvan todas estas cuestiones, voy a resumir la teoría y a presentar las predicciones que se desprenden de ella.

En las ciencias afectivas es ampliamente aceptado que las emociones son estados compuestos por múltiples aspectos. Por ejemplo, Scarantino y De Sousa mencionan que en un episodio de miedo intenso cuando de pronto aparece un oso en el camino de alguien, dicho episodio involucra diferentes componentes, entre ellos: "*un componente evaluativo* (ej. evaluar el oso como peligroso), *un componente fisiológico* (ej. aumento de la frecuencia cardiaca y presión arterial), *un componente fenoménico* (ej. una sensación desagradable), *un componente expresivo* (ej. párpados levantados, mandíbula abierta y caída, labios contraídos), *un componente conductual* (ej. una tendencia a huir) y *un componente mental* (ej. enfocar la atención)" [Scarantino y De Sousa, 2018].

De manera general, la sensación desagradable en el caso del miedo, refiere a la *valencia* del estado. Esto es, el aspecto positivo (agradable) o negativo (desagradable) de la emoción. Por otro lado, el componente fisiológico refiere a la *excitación/activación* fisiológica o corporal.

Ahora bien, Colombetti y Kuppens (2025) han señalado que estos términos suelen ser ambiguos, pues a menudo se utilizan para referir a diferentes aspectos de las emociones. Respecto al término "valencia", sostienen que en ocasiones se utiliza para indicar si una experiencia se siente agradable o desagradable/placentera o no placentera; en otras ocasiones se utiliza para indicar conductas de aproximación (valencia positiva) o de retirada (valencia negativa); también se puede utilizar en relación a las evaluaciones que un sujeto puede hacer acerca de una situación u objeto (evaluar X como positivo o como negativo) y, por último, puede referir a un objeto o a un estímulo (el objeto o estímulo en sí tiene una valencia negativa o positiva).

Por otro lado, sostienen que "activación" o "excitación" son términos que pueden utilizarse para indicar sensaciones de activación que van desde bajas (cuando te sientes

cansado, por ejemplo) a altas (cuando sientes que tienes mucha energía o te sientes muy ansioso). Pero también se pueden utilizar para indicar cambios fisiológicos o corporales que pueden sentirse o no. Por ejemplo, cambios en la frecuencia cardiaca, la respiración, la sudoración, el flujo sanguíneo, etc[46].

Sin embargo, de utilizarse estos términos en todos estos sentidos, podría pensarse que las emociones solo están compuestas de valencia y activación. Esto no parece ser el caso. Precisamente por esta razón, Scarantino y De Sousa distinguen valencia (identificada únicamente a nivel fenoménico o experiencial) de otros aspectos que no la involucran, pero que están relacionados tanto con ella, como con la activación fisiológica que típicamente se siente en un episodio de miedo.

Por su parte, Russell (2003) distingue diferentes tipos de valencia, aunque pueden estar íntimamente conectados. Por ejemplo, de manera general, afirma que alguien puede sentirse feliz (y, de acuerdo al autor, la felicidad es una emoción que *se siente* agradable) porque percibe un objeto o estímulo agradable (cabe señalar que, para el autor, la valencia es intrínseca a los objetos o estímulos). Es decir, hay una experiencia emocional positiva que tiene su causa en un objeto que tiene una valencia positiva. No obstante, advierte que esto no siempre es el caso. Alguien puede percibir un objeto con valencia positiva y no experimentar ninguna emoción.

Más aún y relacionado con este punto, sostiene que, aunque los estados afectivos incorporan dimensiones afectivas como la valencia y la activación, no se reducen a ellas. También incorporan otros elementos como las *evaluaciones* de las situaciones y "la búsqueda de *opciones conductuales* que maximicen el placer y minimicen el displacer" [Russell, 2003, p. 149. Énfasis mío]. Luego, distingue diferentes tipos de valencia que pueden estar relacionados y los distingue, al igual que Scarantino y De Sousa, de otros aspectos de las emociones.

[46] Estos términos son menos problemáticos, pues aunque podría ser difícil explicar cuáles son los cambios fisiológicos que siente un sujeto cuando está cansado o tiene energía, hay otros estados que indican una activación fisiológica alta. Por ejemplo, cuando alguien se siente ansioso, lo identifica por la taquicardia que siente.

A partir de todo lo anterior y para evitar en la medida de lo posible controversias respecto a estos términos, voy a adoptar una concepción de la valencia según la cual es el aspecto positivo o negativo de (o atribuido a) un evento, objeto, estímulo o estado mental y, que puede estar relacionado con otros aspectos de las emociones, como el aspecto conductual y/o motivacional. Esto implica que me mantengo neutral respecto a si la valencia es *en sí* parte de un objeto, estímulo o situación o si percibimos estos objetos y situaciones *como* positivos o negativos debido a que los evaluamos como tal.

Respecto a la activación, voy a entenderla como activación fisiológica o corporal que *típicamente* el sujeto nota o siente. Es decir, normalmente el sujeto siente los cambios corporales, como el aumento en la frecuencia cardiaca, el aumento de la presión arterial, que respira más agitado en ciertas circunstancias, etc.

Si bien es cierto que las emociones son estados compuestos por dimensiones afectivas como la valencia y la activación (entre otros aspectos), como he señalado previamente, se ha postulado que otros tipos de estados mentales también las incorporan. Siendo así, no voy a tomar estas dimensiones afectivas como aspectos exclusivos de las emociones.

Comencemos con las sensaciones corporales. Estos estados también son estados compuestos que no solo incluyen una sensación en alguna parte del cuerpo, sino también una valencia, una activación fisiológica o corporal, así como aspectos motivacionales (véase Aydede, 2019; Pitcher, 1970; Armstrong, 1962 y Tye, 2006).

Consideremos los dolores. Estos tienen (o suelen tener) una valencia hedónica negativa. Esto es, se sienten desagradables; cuando el dolor es intenso, por ejemplo, hay contracción muscular, sudoración y cambios en la respiración. Luego, hay activación fisiológica o corporal. Por último, si sientes dolor, lo natural es que quieras que pare, así que hay una tendencia a retirarse o a alejarse del dolor o de aquello que lo causa. Esto último corresponde al aspecto motivacional[47].

[47] Hay un debate respecto a la relación entre la valencia y el aspecto motivacional que puede capturarse a partir del contenido de la valencia. Por un lado, Carruthers (2018) sostiene que la valencia tiene un *contenido evaluativo*. Es decir, la valencia representa una sensación, por ejemplo*, como* mala (valencia negativa) o *como* buena (valencia positiva). En el caso del dolor, afirma que la experiencia representa una condición corporal (una sensación punzante en el pie, por ejemplo) como mala. Más aún, sostiene que la valencia es directamente motivadora. Si la sensación corporal se representa o se evalúa como mala,

Otros estados que pueden incluir o incorporar dimensiones afectivas son los pensamientos y, como sugiere Russell, las percepciones en general[48]. Es claro que, cuando menos algunos pensamientos, son agradables y otros no. Luego, no solo nos percatamos de su contenido, sino que el mismo contenido es (o se considera como) positivo o negativo. De la misma manera, las percepciones suelen[49] tener una valencia. Por ejemplo, una comida, un sonido o una aroma pueden ser agradables o desagradables. Más aún, si una percepción es agradable o desagradable, puede influir o reflejarse en la activación fisiológica. Por ejemplo, si de pronto percibo un ruido muy fuerte, no solo será una percepción desagradable para mí, sino que mi corazón comenzará a latir más rápido.

En resumen y, a partir de lo anterior, la teoría afectivista plantea que múltiples estados mentales incorporan dimensiones afectivas específicas como la valencia y la activación fisiológica. La valencia es el aspecto positivo o negativo de (o atribuido a) un evento, objeto, o estado mental. De tal forma que, por ejemplo, se puede percibir un objeto agradable/desagradable o placentero/no placentero (o como agradable/desagradable o como placentero/no placentero). De la misma manera, un estado específico, como una emoción, una percepción o un pensamiento, puede sentirse agradable o desagradable. Por otro lado, la activación fisiológica involucra cambios corporales o fisiológicos que típicamente se sienten, como la sudoración, la frecuencia cardiaca, los cambios en la respiración, etc. Más aún, dichas dimensiones afectivas pueden dar lugar a tendencias de acción específicas, como retirarse ante un objeto desagradable o evitar aquello que causa daño.

---

entonces esto directamente motiva a retirarse de aquello que causa dicha sensación.
Por otro lado, Barlassina y Hayward (2019) sostienen que la valencia tiene un *contenido imperativo*. Esto es, una experiencia es agradable o desagradable en virtud de que su contenido ordena al sujeto a obtener más u obtener menos de esa experiencia. En el caso del dolor, Barlassina afirma que, "cuando sientes la sensación desagradable de un dolor de muelas, sientes la inclinación de deshacerte de esa misma experiencia" [2020, p. 277]. En este caso, la valencia es en sí misma motivadora. Luego, podemos pensar que, en la teoría de Carruthers, la motivación es una consecuencia de un paso previo. Esto es, de evaluar X como malo o como bueno. Mientras que, en la teoría de Barlassina y Hayward, la motivación es constitutiva de la valencia y/o la valencia en sí misma tiene fuerza motivacional.
No me comprometo con ninguna de las dos opciones. Lo importante para mi propósito, es que la valencia, de una manera u otra, está relacionada con la motivación.

[48] Otros autores, como Barrett y Lida (2025); Lebrecht et al. (2012) y Satpute et al. (2015), sostienen que no solo estos estados incorporan estas dimensiones afectivas, sino también las creencias y recuerdos. Más aún, defienden que son características *fundamentales* de ellos.

[49] Autores como Zajonc (1980) y Lebrecht et al. (2012) sostienen que *todas* las percepciones tienen una valencia intrínseca o, cuando menos, lo que Lebrecht y sus colegas denominan "micro-valencias".

Así, si alguna de estas dimensiones afectivas (o ambas) está ausente o disminuida de estos estados, entonces hay déficits afectivos que provocan las experiencias anómalas que reportan los pacientes despersonalizados. En otras palabras, las anomalías de las experiencias que expresan estos pacientes son causadas por déficits afectivos que consisten en ausencia o disminución de valencia o activación fisiológica.

Ahora bien, en los dos capítulos anteriores vimos que, tanto la teoría gradualista como la negacionista, sugieren que debemos tomar los reportes de manera literal y/o que, para tomarnos seriamente lo que dicen los pacientes, debemos interpretar sus reportes literalmente. De otra manera, no podríamos dar sentido a sus discursos.

La teoría afectivista no sigue esta sugerencia o máxima metodológica de estas dos teorías. Sin embargo, esto no implica que no se tome en serio los reportes de los pacientes o que no pueda dar sentido a lo que expresan. Pensemos que, en el habla cotidiana, es común recurrir a expresiones metafóricas o a analogías para intentar describir cómo nos sentimos o cómo nos parece que son las cosas, pues, en muchas ocasiones, es difícil explicar lo que sentimos recurriendo al lenguaje literal. Por ejemplo, solemos decir que sentimos mariposas en el estómago cuando alguien nos gusta. Claramente la descripción no es literal, pero esto no significa que no tenga sentido lo que estemos diciendo o que sea incomprensible. Tampoco significa que no nos tomemos en serio lo que estamos sintiendo (o que alguien más no se tome en serio lo que le están expresando[50]).

Siendo así, voy a recurrir al habla cotidiana en la que, como he señalado, la gente común (no filósofa o psicóloga) utiliza analogías o metáforas para describir cómo se siente. Luego, no debemos suponer que los pacientes conocen los términos a los que recurre la teoría afectivista y, que en consecuencia, reportarán literalmente que sus estados carecen de valencia o activación. Incluso si literalmente reportan ausencia de emociones, es poco

[50] De manera similar, en el estudio de la consciencia es común recurrir a metáforas para da cuenta de aspectos de la consciencia que son difíciles de explicar recurriendo al lenguaje literal. Pensemos, por ejemplo, en la expresión "pintura mental" (Block, 2003) que trata de dar cuenta de un aspecto no intencional de la consciencia. O en expresiones como "X está *en el flujo* de consciencia de tal sujeto", que pretende indicar que X le pertenece al sujeto (véase Billon, 2013, quien menciona diferentes metáforas espaciales que tratan de indicar un sentido de pertenencia). En estos casos, no se nos está invitando a que no nos tomemos en serio lo que se está diciendo. Antes bien, son recursos habituales para describir aspectos sumamente elusivos, para los cuales no parece estar diseñado el lenguaje natural.

probable que sepan que éstas son estados compuestos y, mucho menos, por estas características específicas.

Sin embargo, aunque no se expresen de manera literal o aunque no utilicen los términos teóricos que caracterizan a la teoría afectivista, voy a mostrar que sus reportes apuntan a las dimensiones que se han planteado. Regresemos al ejemplo de las mariposas en el estómago. Si bien la expresión no es literal, puede formularse *teóricamente* en términos de la activación fisiológica que siente un sujeto (aunque, de nuevo, esto no significa que el sujeto sabe qué es esto). Otro ejemplo puede ser aquel en el que solemos decir que una comida no nos sabe a nada cuando no la estamos disfrutando o cuando no es placentera (aunque sabemos si algo es placentero o no, no sabemos que esto corresponde a la valencia del objeto, por ejemplo). Luego, la expresión "la comida no sabe a nada" no es literal, pero, de nuevo, *teóricamente* podemos explicarla en términos de una falta o ausencia de algo que, normalmente está, pero que puede estar ausente.

Esta estrategia a la que voy a recurrir tiene respaldo en algunas descripciones de la despersonalización y en algunos reportes de los pacientes. Hemos visto que algunos sujetos reportan que sienten como si estuvieran muertos, lo que parece ser una analogía o metáfora[51]. Por otro lado, Mayer-Gross (1935) señaló lo siguiente respecto a las descripciones de los pacientes despersonalizados: "lo que realmente ha cambiado o disminuido con la aparición de la despersonalización no puede expresarse en el discurso. Incluso personas educadas (como en algunos casos en la literatura) no han dado una descripción clara, sólo han utilizado metáforas" [1935, p. 166].

De manera similar, Sierra y David (2011) han señalado que los pacientes se percatan de que hay un cambio en sus experiencias aunque este es inefable, por esta razón, recurren al uso de metáforas, como estar aislado o separado del mundo "por una barrera invisible como un cristal, niebla o velo" [2011, p. 99].

En vista de lo anterior, la teoría afectivista no predice un tipo de reporte en específico. Sin embargo, la teoría afectivista puede predecir que, si hay déficits afectivos, los pacientes reportarán ausencia o disminución de emociones. Alternativamente, puede predecir que, si

[51] Pensemos que cuando nos sentimos muy mal físicamente, por ejemplo, solemos decir que sentimos que nos estamos muriendo. Esto no es literal, sino figurativo.

diferentes estados están *compuestos* y/o involucran dimensiones afectivas, al haber disminución o ausencia de dichas dimensiones, los pacientes notarán que algo que normalmente estaba presente ya no está o parece haberse disminuido. En consecuencia, utilizarán expresiones no literales para tratar de explicar esta ausencia o disminución.

En la siguiente sección, voy a interpretar los reportes de los pacientes a partir de esta teoría y sus predicciones y voy a ir señalando cómo evita o cómo puede resolver los problemas a los que se enfrentaban las teorías pasadas. Dado que, como hemos visto, los reportes son de diversos tipos, voy a dividirlos por estados. Primero, voy a interpretar reportes que refieren a emociones; después, reportes que refieren a dolores; posteriormente, reportes que refieren a estados cognitivos y perceptivos. Por último, voy a interpretar reportes que no refieren a un estado en particular y aquellos que, en primera instancia, no parecen apuntar a ningún estado mental.

3.4. La interpretación afectivista de los reportes

Reportes que refieren a emociones: Comencemos por los reportes que apelan a emociones o que sugieren algún problema con las emociones. En capítulos anteriores encontrábamos los siguientes: "besar a mi esposo es como besar una mesa, señor. La misma cosa… ni la menor emoción. Nada en la tierra puede emocionarme. Ni mi esposo ni mi hijo…mi corazón no late. No puedo sentir nada"; "un estado de nada, sin ningún estado de ánimo, como si estuviera muerto"; "parece que estoy caminando en un mundo que reconozco pero que no siento"; "puedo sentirme entumecido de sentimientos, casi vacío por dentro. Odio el hecho de que no puedo sentir las cosas como solía hacerlo" y "la parte emocional de mi cerebro está muerta… todas mis emociones están aplanadas".

En el primer caso, la mujer está reportando ausencia de emociones, tanto positivas como negativas. Esto es, no siente amor ni odio cuando besa a su esposo y/o no le produce placer ni displacer besarlo. En general, nada de lo que sucede a su alrededor le provoca una emoción. Luego, hay ausencia de emociones en general. Veíamos que la teoría gradualista suponía que, cuando la subjetividad se atenúa de las emociones, el sujeto podría sentir como si no tuviera emociones *en absoluto*, pero, señalé que esto no es coherente con la predicción de la teoría, pues esta predice que, si la subjetividad se atenúa de los estados, los sujetos reportarán *como si no fueran suyos*, no como si no los tuvieran. Por otro lado, la mujer

tampoco expresa que es *como si* no tuviera emociones. Ya vimos que reporta ausencia total de ellas y, si hay ausencia, entonces no hay o no existe un estado emocional que pueda ser subjetivo ni en menor ni en mayor grado.

Por otra parte, la teoría gradualista no dice nada respecto a la parte del reporte donde la paciente expresa que su corazón no late. La teoría afectivista tiene una ventaja adicional ante este reporte, pues no solo muestra que hay ausencia de estados emocionales tanto positivos como negativos, sino que puede interpretar esta última expresión. Por ejemplo, es común que cuando vemos a una persona que queremos, nuestro corazón late más fuerte. Esto es, en términos teóricos, hay aumento de la activación fisiológica. Ahora, es posible que lo que trata de expresar cuando dice que su corazón no late, es que no siente que comienza a latir *más fuerte* cuando ve a su esposo o a su hijo, por ejemplo. Es decir, no quiere decir literalmente que su corazón ha dejado de latir, simplemente ya no siente cómo se acelera su corazón en presencia de sus seres queridos[52]. Luego, la expresión "mi corazón no late" parece ser la más adecuada para describir la ausencia de esta sensación y para reforzar la incapacidad que tiene para experimentar emociones.

En el segundo reporte también encontramos ausencia de estados afectivos, pues precisamente reporta que no tiene *ningún* estado de ánimo. Si hay ausencia, entonces en este caso tampoco existe un estado afectivo que pueda ser subjetivo ni en menor ni en mayor grado. Por otro lado, como sugerí desde el primer capítulo, la referencia a la muerte no es literal, es una metáfora que utiliza para expresar precisamente la deficiencia que tiene para experimentar estados de índole afectivo. Ahora bien, si restringimos el reporte a un déficit afectivo, como de hecho sugiere el mismo sujeto que lo expresa, evitamos recurrir a los supuestos implausibles que plantea la teoría gradualista, como suponer que el sujeto se refiere

[52] Si bien en este caso no se da la activación fisiológica y/o si no siente los cambios fisiológicos, otros pacientes reportan que es lo único que pueden sentir. Por ejemplo, un sujeto reporta que solo siente ira por fuera, por sus reacciones fisiológicas (véase Apéndice, reporte 27). En este caso, quizás siente calor en la cara, así como los puños y la mandíbula contraídos. Esto sugiere que el sujeto identifica las reacciones típicas de la ira, así que la sigue considerando como tal. Es decir, como un episodio de ira. Sin embargo, sugiero que en este caso falta otro aspecto que compone dicha emoción. Si sentir ira es sumamente desagradable o una emoción que se siente desagradable, al faltar esta valencia negativa típica de la emoción, al sujeto le puede parecer que solo siente ira por fuera.

a que se siente *fenoménicamente muerto* porque la subjetividad se ha atenuado de *la mayoría* de sus estados mentales.

Respecto al reporte del paciente que camina por un mundo que reconoce pero que no siente, vimos que la teoría gradualista no especificaba, en primer lugar, por qué este reporte involucraba una emoción y, aunado a esto, tampoco especificaba qué podría querer decir el paciente cuando afirma que reconoce el mundo, pero no lo siente. Y, como señalé, una buena teoría debería ser capaz de indicar cómo se relacionan las emociones con reconocer el mundo, pero no sentirlo.

La teoría afectivista puede resolver estas cuestiones. Primero, parece plausible que en el habla cotidiana cuando alguien dice que no siente el mundo, puede querer decir que *no siente nada* cuando ve ciertas escenas u objetos del mundo que lo rodean. Esto puede interpretarse de dos maneras: 1) al decir esto, el sujeto puede referirse a que no hay nada que se sienta para él ver dichas escenas u objetos. Esto es, puede referirse a que no tiene un estado visual fenoménicamente consciente. Sin embargo, hemos visto en los capítulos anteriores que es poco plausible que un sujeto ordinario tenga el conocimiento de qué es un estado fenoménicamente consciente y uno inconsciente. 2) al decir esto, el sujeto puede referirse a que dichas escenas u objetos del mundo *no le provocan ninguna emoción*. Esta opción resulta más plausible en tanto, comúnmente, sentimos emociones cuando vemos, por ejemplo, un cuadro que nos gusta o escuchamos una canción que puede tener un significado especial para nosotros.

Por otro lado, en la sección 3.2, señalé que Gerrans (2015) o Sierra y David (2011) han argumentado que estos sujetos exhiben ausencia de respuestas afectivas o de respuestas emocionales, no solo en relación con sus seres queridos, sino también en relación con situaciones u objetos del mundo que normalmente las causaban. Para ilustrar este punto, Sierra y David recurren al siguiente reporte extraído del trabajo de Bockner (1949): "[al escuchar música] no hay respuesta en mí. La música normalmente *me conmueve*, pero ahora bien podría ser alguien picando patatas (…) anoche vi el Big Ben iluminado, normalmente

una imagen conmovedora para mí, pero podría haber sido un despertador por lo que sentí"[53] [Bockner, 1949, p. 969. Énfasis mío]

En este caso, claramente ya no se produce una emoción en el sujeto o no hay una respuesta emocional al escuchar música o al ver el Big Ben. Luego, de manera similar, el sujeto que reporta que camina por un mundo que reconoce, pero que no siente, puede referirse a que ve los objetos y escenas del mundo que lo rodean, pero estos ya no lo hacen emocionarse como solía suceder antes del episodio de despersonalización o que ya no le provocan ninguna emoción.

Por último, los últimos dos reportes refieren a emociones "entumecidas" o "aplanadas". En la sub-sección 1.4.2 del primer capítulo, señalé que el entumecimiento o el aplanamiento podían entenderse como una reducción en la intensidad de las emociones. También señalé que, si se entendía de esta manera, la teoría gradualista tenía que recurrir a una hipótesis adicional según la cual si la subjetividad de la emoción disminuía, entonces la emoción se veía disminuida en intensidad. Sin embargo, no ofrecía una razón independiente que mostrara que esto podía suceder. Además, vimos que, al parecer, el sujeto podía sentir *sus* emociones, aunque estas fueran menos intensas. Luego, no parecía haber una relación entre la subjetividad disminuida y la disminución de la intensidad de una emoción.

Por otro lado, en la sección 2.3 del segundo capítulo, señalé que la teoría negacionista no disponía de recursos para explicar una emoción disminuida pues es una alternativa a la gradualista y, aunque incorporara aspectos de dicha teoría, se enfrentaba a las mismas dificultades que ella. Siendo así, ambas teorías presentaban dificultades para interpretar este tipo de reportes.

[53] Cabe señalar que, a partir de este reporte, los autores mencionan lo siguiente: "tales afirmaciones parecen sugerir que lo que se ve más afectado en la despersonalización es la capacidad de impregnar a los objetos percibidos o a las situaciones concretas de sentimientos emocionales, en lugar de una incapacidad general para experimentar estados emocionales" [2011, p. 101]. Sin embargo, los autores nunca especifican qué significa "impregnar objetos de sentimientos emocionales". Parece natural pensar que los sentimientos emocionales *los siente* un sujeto al ver dichos objetos. De ser así, si no hay sentimientos emocionales hacia tales objetos, entonces el sujeto no los experimenta. Luego, no es claro cómo difiere esto de la incapacidad de experimentar estados emocionales. Más aún, este caso precisamente señala que el sujeto tiene problemas para sentir emociones.

La teoría afectivista puede dar cuenta de la reducción de la intensidad de una emoción al apelar a la dimensión afectiva de la activación fisiológica. Por ejemplo, mencionaba que, cuando vemos a alguien que queremos o experimentamos amor por alguien, típicamente aumenta nuestra frecuencia cardiaca (nuestro corazón late más fuerte) y solemos sentir algo en el estómago que muchas veces describimos como "mariposas en el estómago" (más aún, dichas sensaciones suelen ser agradables). Por otro lado, cuando tenemos miedo, sentimos que nuestro cuerpo (o alguna de nuestras partes corporales) tiembla, sentimos cambios en la piel que solemos describir como "piel de gallina" y una sensación de escalofríos en la espalda (y dichas sensaciones suelen ser desagradables)[54].

Ahora, si en estos pacientes disminuye la intensidad de la activación fisiológica (la frecuencia cardiaca cuando ven a un ser amado, la sensación de escalofríos en la espalda o incluso si todos estos cambios a nivel fisiológico disminuyen al mismo tiempo), entonces podrían decir que sus emociones están disminuidas, aplanadas o entumidas. Luego, parece haber una relación entre una emoción disminuida y una activación fisiológica reducida[55]. Como señalé en la sección 3.3, incluso si en estos casos los sujetos reportan literalmente que sus emociones están disminuidas (debido a la reducción de la activación fisiológica), es poco probable que sepan que, precisamente se sienten disminuidas por aquello que *teóricamente* llamamos activación fisiológica. Un sujeto común simplemente siente disminuidos ciertos cambios corporales.

Reportes que refieren a dolores: En la sección 3.3, mencioné que es ampliamente aceptado que los dolores también incorporan dimensiones afectivas como valencia y activación fisiológica y que, al igual que las emociones, también incorporan aspectos motivacionales.

[54] Respecto al miedo, hay pacientes que pueden expresar que, aunque tienen miedo, este no es genuino (véase Apéndice, reporte 28). Si bien en este reporte no encontramos grados, a partir de la identificación de las reacciones fisiológicas del miedo, podemos pensar que no le parece un miedo genuino porque no se dan estos cambios fisiológicos y/o porque no los siente. Alternativamente, puede sentir, por ejemplo, la sudoración fría y los escalofríos en la espalda y no sentirlos como desagradables.

[55] Probablemente uno de los casos que mejor puede ilustrar este punto es el de la ansiedad. Si el lector alguna vez en su vida ha experimentado ansiedad, sabrá que esta se identifica inmediatamente por la taquicardia que se siente y por la repentina sudoración excesiva en las manos (si bien no son síntomas exclusivos, son síntomas que aparecen de manera inmediata y fácilmente reconocibles). Ahora, hay técnicas o medicamentos para controlarla y, cuando esto sucede, aunque persiste la taquicardia y la sudoración, estos cambios fisiológicos disminuyen y, en consecuencia, sientes cómo la ansiedad va disminuyendo o aplanándose.

En la sección 2.3 del segundo capítulo, vimos que hay pacientes que reportan que se sienten desapegados de sus dolores o que estos no les parecen genuinos. También vimos que la teoría negacionista no podía dar cuenta de estos reportes a menos que añadiera una hipótesis adicional. En términos generales, podía suponer que, aunque los pacientes no reportaran explícitamente que sus dolores no se sentían como si fueran de ellos, de cualquier manera no eran subjetivos y, por tanto, eran fenoménicamente inconscientes y por esta razón podían sentirse desapegados de ellos o sentir que no eran genuinos. Sin embargo, el supuesto resultaba poco plausible debido a que los mismos reportes sugerían que los sujetos podían *sentir* sus dolores. Luego, todo apuntaba a que había fenomenología.

La teoría afectivista no necesita recurrir a hipótesis adicionales que resultan ser poco plausibles, pues si como acabo de señalar, los dolores son *estados compuestos* por dimensiones afectivas (entre otros aspectos), entonces es posible que la ausencia o la disminución de dichas dimensiones pueda dar cuenta de las experiencias de dolor que reportan los pacientes.

Comencemos por el reporte del paciente que expresa desapego: "cuando una parte de mi cuerpo me duele, me siento tan desapegado del dolor que se siente como si fuera el dolor de alguien más". Mencionaba que el desapego se suele definir como falta de interés o indiferencia hacia algo. Ahora, es plausible que si al estado de dolor le falta la valencia hedónica típica de él, entonces el sujeto puede sentir dolor, pero ya no lo siente desagradable y, si ya no se siente de esta manera, entonces el dolor no le molesta al sujeto o le es indiferente. Para mostrar que esto puede suceder, pensemos en los casos de administración de morfina. Los sujetos suelen reportar que cuando se les administra, les sigue doliendo una parte corporal, por ejemplo, pero que ya no les molesta o que les da igual, precisamente porque falta el aspecto negativo del dolor (véase Aydede, 2019 y Morales, 2023)[56].

[56] También señalan que lo mismo sucede con una condición conocida como "asimbolia del dolor". Por ejemplo, Aydede menciona que, "las personas que la padecen experimentan dolor sin el afecto negativo. Curiosamente, estas personas aún identifican su experiencia como dolor, pero no muestran signos corporales, emocionales ni conductuales típicamente asociados con el aspecto desagradable del dolor" [Aydede, 2019]. De manera similar, Morales señala que, tanto en los casos de administración de morfina, como en la asimbolia del dolor, los pacientes no solo reportan que no sienten el aspecto desagradable del dolor, sino que no se sienten inclinados a evitarlo. "Esto no significa que encuentren el dolor placentero, sino que simplemente les es indiferente y no se sienten impulsados a evitarlo" [Morales, 2023, p. 4].

Ahora bien, si el dolor ya no es molesto debido a que falta la valencia negativa de ese estado, entonces también es plausible que desparezca el aspecto motivacional asociado con la valencia, pues, como vimos en la sección 3.3, de acuerdo a Carruthers (2018), una sensación corporal evaluada *como mala* (valencia negativa) motiva directamente a retirarse de aquello que causa la sensación mala o desagradable. O, de acuerdo a Barlassina y Hayward (2019) y Barlassina (2020), si hay una sensación desagradable en una parte corporal, entonces se siente la inclinación de deshacerse de la experiencia de dolor.

Como señalé respecto a estas dos posturas, no es necesario comprometernos con una en particular, pues, si en la teoría de Carruthers, la motivación es una consecuencia de la evaluación de una sensación corporal como mala, si no se evalúa como tal, entonces no hay motivación para retirarse de aquello que causa malestar o desagrado. O, si en la teoría de Barlassina y Hayward, la valencia en sí misma es motivacional y/o tiene fuerza motivacional, al faltar la valencia, no se siente la inclinación de deshacerse de la experiencia desagradable de dolor.

Por otro lado, al decir que es como si fuera el dolor de alguien más, no tenemos buenas razones para pensar que, de hecho, se lo están atribuyendo a alguien más, pues ya vimos que no solo no hay errores de auto-atribución, sino que el mismo reporte sugiere que es él mismo quién lo siente. Siendo así, podríamos suponer que el sujeto está utilizando una analogía. Por ejemplo, el dolor le puede ser indiferente al sujeto en tanto no le molesta (así como en los casos de administración de morfina) y, así como no le molestan sus propios dolores, es *como si* fuera el dolor de alguien más en tanto el dolor ajeno de alguien no lo molesta a él. Si bien es un supuesto adicional, resulta plausible en tanto uno no puede sentir la molestia de un dolor ajeno.

Veamos ahora el reporte del dolor que no parece genuino: “era doloroso y sentí que mi brazo se retraía, pero no era un dolor genuino, era un dolor que no llegaba al alma… es un dolor, si quieres, pero la superficie de mi piel está tres kilómetros lejos de mi cerebro y no sé si estoy sufriendo.” Al igual que con el caso anterior, si los dolores son estados compuestos, si falta alguno de sus componentes es probable que no parezca un dolor genuino o real. De nuevo, pensemos que si son desagradables y si también nos causan sufrimiento (que, si bien no es una emoción, puede ser un estado de ánimo que suele ser prolongado e

involucrar no solo dolor, sino angustia. Luego, es una experiencia afectiva), si falta la valencia negativa y si no provoca una experiencia afectiva como el sufrimiento, entonces puede parecerle al sujeto que no es un dolor genuino o real.

Adicionalmente, parece que en este caso hay otros componentes que están ausentes o disminuidos. Por ejemplo, el sujeto está sugiriendo que su dolor no era tan intenso como solía serlo antes de su episodio de despersonalización. Esto podemos apreciarlo con las metáforas que utiliza: "era un dolor que no llegaba al alma" o "la superficie de mi piel está tres kilómetros lejos de mi cerebro". Ahora bien, la intensidad del dolor también podemos explicarla al apelar a la activación fisiológica. Como señalé anteriormente, un dolor intenso puede caracterizarse por la contracción muscular, sudoración en todo el cuerpo y cambios en la respiración (puede costarte trabajo respirar o puedes respirar agitadamente). Si la activación fisiológica se da en un menor grado o, incluso si se da, pero el sujeto no siente estos cambios fisiológicos, es probable que exprese que el dolor no "llega al alma" (no es intenso como solía ser un dolor del tipo que está experimentando). Por otro lado, parece que también hay ausencia de un componente motivacional, pues expresa que sentía que su brazo se retraía, pero no indica que él mismo tuviera la disposición o la inclinación de retirarlo para calmar el dolor o la inclinación para deshacerse de la experiencia desagradable del dolor. Si bien la ausencia del componente motivacional puede explicarse por la ausencia del aspecto desagradable (valencia negativa), también es posible que influyan otros factores, como la intensidad disminuida y porque el dolor no le causa sufrimiento al sujeto.

Esta interpretación de los reportes no solo tiene la ventaja de no apelar a supuestos poco plausibles (o incluso que van en contravía de lo que sugiere la teoría), sino que muestra que hay un tipo de fenomenología (sensorial), pero no otro tipo (afectiva). De esto se desprende, a diferencia de lo que plantea la teoría negacionista, que hay fenomenología parcial.

Reportes que refieren a estados cognitivos y perceptivos: He señalado que los pensamientos, creencias y percepciones también suelen incorporar una valencia positiva o negativa. Ahora, si estos estados normalmente estaban valenciados o si incorporaban esta dimensión y, si en algún momento la pierden, es probable que el sujeto lo note (de otra

manera no notaría un cambio en la experiencia) y que trate de explicar este cambio recurriendo también a expresiones no literales.

Consideremos los pensamientos. Veíamos en el segundo capítulo (sección 2.3) que un paciente reporta lo siguiente: “me siento tan desapegado de mis pensamientos que parecen tener una ‘vida’ propia.” En este caso, los negacionistas podían suponer que el paciente se siente despegado de los pensamientos porque, aunque no exprese explícitamente que no se sienten como si fueran de él, de hecho no los siente propios y por esta razón se siente desapegado de ellos. No obstante, claramente el sujeto está expresando que se siente desapegado de *sus* pensamientos. Luego, esto no resulta plausible. Otra opción que tenían era suponer que mientras *sabe* que son de él, no los *experimenta* de esa forma y por dicha razón expresa que parecen tener vida propia.

Claramente esta expresión no es literal, precisamente por esta razón puede entenderse de diferentes maneras. Si bien es cierto que puede señalar que el sujeto *no siente* que son sus pensamientos, aunque sabe que son de él, mencionaba que cuando los pensamientos surgen recurrentemente sin ser invitados o sin que queramos pensar en algo en específico, una expresión adecuada es que parecen tener vida propia.

Ahora, cuando de pronto surgen pensamientos que no queremos pensar, típicamente son pensamientos negativos. Luego, es posible que al paciente no le interesen o sienta indiferencia hacia dichos pensamientos, precisamente porque tienen una valencia negativa (nadie se siente particularmente atraído o interesado por sus pensamientos negativos, precisamente uno trata de evitarlos). Sin embargo, si siguen apareciendo recurrentemente, podría decir que parecen tener una vida propia. Por supuesto que en este caso no habría ausencia de valencia, el problema estaría en que solo hay pensamientos con valencia negativa. No obstante, podríamos pensar que hay un déficit afectivo que consiste en que sólo hay pensamientos negativos.

Otra opción podría ser que ninguno de los pensamientos que aparecen recurrentemente le interesen porque ninguno es especialmente agradable/desagradable o atractivo/repulsivo. Luego, mientras siguen apareciendo pensamientos en su cabeza, ninguno logra atraer su interés. Esto no implica que el paciente no sienta que son sus pensamientos.

Aunque puede sentir que son suyos, le son indiferentes o no le interesan porque, o solo son negativos o porque todos carecen de valencia negativa y positiva.

Luego, a diferencia de lo que podría suponer la teoría negacionista, el hecho de que reporte que parece que sus pensamientos tienen vida propia, no implica *necesariamente* y/o *exclusivamente* que no los experimenta como suyos.

Consideremos ahora las percepciones. Veíamos en el segundo capítulo que algunos sujetos reportan que todo en la visión está muerto o que, aunque pueden ver, oír y oler, es como si no vieran, oyeran u olieran. Por el lado de la fenomenología, la teoría negacionista planteaba que, en estos casos, los pacientes sugerían que no hay nada que se sienta para ellos ver, oír u oler. Luego, que dichos estados no son fenoménicamente conscientes. Sin embargo, a partir del caso hipotético de la súper visión ciega, señalé también que si los estados que reportan los sujetos no se sintieran de cierta manera o si no hubiera nada que se sienta para los sujetos ver, oír u oler, entonces cabría esperar que reportaran que no pueden ver, oír u oler. Pero esto no es lo que encontramos. Por tanto, en principio parece que estos estados son fenoménicamente conscientes.

Si lo anterior es el caso, entonces podríamos pensar que es otra característica o cualidad de la experiencia la que está ausente. Veamos cómo se interpretan estos reportes bajo el marco de la teoría afectivista. De nuevo, uno de los sujetos reporta que todo en la visión está muerto. Si hasta donde sabemos el sujeto puede ver, entonces es plausible que cuando afirma esto, se refiere a que los objetos o escenas *que ve*, le parece que están muertos.

La cuestión por resolver es a qué se refiere específicamente por "muertos". Es poco probable que piense que no están vivos así como lo está una persona, un animal o una planta. Lo más probable es que quiere decir que no parecen reales o que parecen extraños. De ser así, el sujeto no está utilizando una expresión literal para señalar cómo le parece que son las cosas.

Ahora bien, la sugerencia de que el sujeto se refiere a que los objetos o escenas del mundo que ve parecen irreales o extraños, es respaldada por la misma descripción de la desrealización. Como hemos visto, en el DSM-V se define como un trastorno disociativo que consiste en experiencias o sensaciones de irrealidad o distanciamiento respecto al mundo y/o

los objetos del mundo y que los pacientes pueden expresar que "el entorno puede verse como artificial, sin color o sin vida" [Ibíd., p. 303].

Por otro lado, en la sección 3.2, señalé que, autores como Gerrans (2015) han argumentado que la sensación de irrealidad que expresan los pacientes despersonalizados se debe a una falta de respuesta afectiva al mundo. Así, señala que "el fracaso de los acontecimientos emocionalmente relevantes para desencadenar respuestas afectivas en el paciente, hace que el mundo se sienta irreal o extraño" [2015, p. 19].

De manera similar, mencioné que Sierra y Berrios (1998) y Sierra et al. (2002) han argumentado que la sensación de irrealidad del mundo se debe a que la percepción del mundo no se "colorea emocionalmente" y/o que la percepción no está dotada de "sentimientos emocionales". Por ejemplo, Sierra et al. (2002) sugieren que los reportes de los pacientes que padecen de hipoemocionalidad visual (no hay respuesta afectiva ante estímulos visuales o ésta está reducida) son similares a los reportes de los pacientes despersonalizados. Para ilustrar su punto, recurren al siguiente reporte de un paciente con hipoemocionalidad visual: "para mí, las flores han perdido su esencia; no las veo como parte de la naturaleza. Se han vuelto casi sintéticas, *artificiales*. Me falta cierto conocimiento; no, no es realmente conocimiento, sino cierta claridad para ver la naturaleza misma. No logro ver la flor en toda su *autenticidad*" [2002, p. 530. Énfasis mío].

Debido a esta similitud con los reportes de los pacientes con despersonalización, concluyen que en ambos casos hay una incapacidad para colorear emocionalmente la percepción y/o para dotar la percepción de sentimientos emocionales y que esto resulta en un cambio en su experiencia perceptiva, que se expresa precisamente como irreal o artificial.

No obstante, como señalé anteriormente (véase nota al pie 53), Sierra y sus colegas no explican qué significa que la percepción o los objetos de la percepción estén dotados (o impregnados) de "sentimientos emocionales". Mencioné que una interpretación natural de esta expresión es que los sentimientos emocionales *los siente* un sujeto al ver dichos objetos. Como en el caso en el que el paciente despersonalizado que citan Sierra y David (2011) ya no se *conmueve* cuando ve el Big Ben o cuando escucha música. Este es un caso que ilustra que el sujeto ya no siente emociones y/o que hay una ausencia de una respuesta emocional al ver objetos que normalmente la provocaban.

Si asumimos esta interpretación natural sobre los sentimientos emocionales, a saber, que el sujeto no siente emociones y/o que hay una falta de respuesta emocional ante los objetos del mundo, entonces no es claro cómo esta falta de respuesta emocional se relaciona con el reporte del paciente con hipoemocionalidad visual y sus reportes de irrealidad, pues no hay nada en el reporte que sugiera que ya no siente una emoción o que hay una ausencia de respuesta emocional, a diferencia del paciente que mencionan Sierra y David (2011).

Sin embargo, esto en sí mismo no significa que debamos de renunciar a una interpretación afectivista. En la sección 3.3, señalé que Russell (2003) sostiene que las percepciones están valenciadas o que incorporan valencias intrínsecas y que, generalmente, la percepción de un objeto con valencia positiva causa una emoción positiva o una respuesta emocional positiva. Por ejemplo, si percibes un atardecer bello[57] o como agradable a la vista, puedes sentirte feliz o sentirte conmovido. Pero, advierte que esto no siempre es el caso. Parece perfectamente posible percibir un objeto agradable o como agradable a la vista y no sentir ninguna emoción o no experimentar ninguna respuesta emocional ante dicho objeto.

Así pues, a partir de las consideraciones de Russell, propongo que, así como el paciente con hipoemocionalidad visual reporta que los objetos que ve le parecen irreales o artificiales, el paciente con despersonalización que reporta que todo en la visión está muerto, también se puede referir a que los objetos o escenas que ve le parecen irreales, extraños o artificiales.

No obstante, a diferencia de lo que afirman Sierra y sus colegas, esto no se debe a que el sujeto no siente emociones y/o a que hay una falta de respuesta emocional ante dichos objetos, sino que se debe a que los objetos no están valenciados o han perdido la dimensión afectiva de la valencia. Así, si el sujeto ya no puede percibir este aspecto positivo de los objetos. Por ejemplo, ya no puede percibir la belleza de las flores o de un atardecer (de nuevo, sin que esto implique que no le provocan una respuesta emocional), entonces le parecerán extraños, como si hubieran perdido una cualidad esencial o como si estuvieran muertos.

[57] Podemos pensar que la belleza de los objetos que percibimos es un aspecto positivo del objeto o que tiene una valencia positiva. De la misma manera, podemos pensar que percibir algo como bello es al mismo tiempo agradable o, una opción alternativa, es equiparar lo agradable con lo bello, como parecen sugerir Shuman et al. (2013). En su propuesta, las evaluaciones de los objetos incluyen la evaluación de algo como agradable/bello o desagradable/feo. Véase también Colombetti y Kuppens (2025).

Pasemos ahora al paciente que reporta que puede ver, oír y oler, pero que es como si no viera, oyera u oliera. La interpretación afectivista en este caso es similar. Si bien el sujeto puede ver, oír y oler, al faltar una cualidad esencial o un aspecto que típicamente está en aquello que percibe, le puede parecer que no percibe *en su totalidad* o que no percibe los objetos *por completo*.

Ya vimos que en el caso anterior, el sujeto reporta que todo en la visión está muerto y que con esto puede referirse a que los objetos parecen irreales o extraños, pero parecen de esta forma, precisamente porque nota que han perdido una cualidad esencial (la belleza de un atardecer, por ejemplo). En este caso, si el sujeto también nota la ausencia de una cualidad esencial (de nuevo, el aspecto positivo o negativo de los objetos), entonces, alternativamente le puede parecer que no puede ver por completo y, en consecuencia, puede decir que es *como si no viera*.

Lo mismo sucedería con las percepciones olfativas y auditivas. En la sección 3.3, ya señalé que los sonidos que escuchamos y los aromas que olemos, también tienen valencias positivas o negativas. Pensemos que el ruido que hace una motocicleta es desagradable y/o lo percibimos como desagradable y hay olores que también percibimos como agradables (un perfume, por ejemplo) o desagradables (el olor a gas). Ahora, si el sujeto ya no puede percibir el aroma del café como agradable, por ejemplo, o si como solemos decir, "ya no huele rico"[58], entonces le puede parecer que falta algo en su percepción y, al faltar algo, le puede parecer como si no oliera (por completo) o, quizás como sugiere el paciente que citan Sierra et al. (2002), como si no lo percibiera *en toda su autenticidad*. Lo mismo con la percepción auditiva, si el sujeto ya no puede percibir la belleza de una melodía que antes consideraba hermosa, entonces le puede parecer que falta algo y que no la percibe por completo o como si no pudiera percibirla totalmente. En general, aunque el paciente puede percibir, al faltar un componente afectivo, le puede parecer como si no lo hiciera.

[58] Estoy considerando valencias positivas, pero también puede haber ausencia de valencias negativas. Por ejemplo, si el paciente antes podía percibir cualquier objeto como repulsivo o desagradable y en su condición actual ya no lo hace, es natural pensar que también le podría parecer que falta algo en la percepción y que es como si no lo percibiera por completo.

Si esto es el caso, entonces al igual que con las sensaciones corporales, de estos reportes se desprende que hay fenomenología parcial. Hay una fenomenología perceptiva, pero no una fenomenología afectiva asociada a ella[59].

Reportes que no especifican un estado en particular: Hasta ahora he considerado reportes que refieren a estados específicos. Sin embargo, hemos visto que, tanto la teoría negacionista como la gradualista, suponen que hay casos en los que *la mayoría* de los estados de los sujetos exhiben subjetividad atenuada y, como consecuencia, pueden reportar que están muertos (refiriéndose a fenoménicamente muertos) o que la mayoría de los estados no son subjetivos, por tanto, no son fenoménicamente conscientes y, como consecuencia, pueden reportar que son conscientes en el sentido de acceso, pero inconscientes en el sentido fenoménico.

Señalé que el problema general con estos supuestos es que es implausible suponer que los sujetos se refieran a que se sienten fenoménicamente muertos (o, alternativamente, como zombies filosóficos) y que se refieran a que son conscientes en el sentido de acceso, pero inconscientes en el sentido fenoménico, porque es poco probable que conozcan estas distinciones o que puedan trazar distinciones finas de la filosofía de la mente que solo están disponibles para un reducido número de expertos.

En vista de esto, un primer paso para dar una mejor interpretación de los reportes, es preguntarnos nuevamente a qué se puede referir la gente *en el habla cotidiana* cuando expresa que se siente muerto o que es consciente e inconsciente. Ya vimos que uno de los sujetos reporta lo siguiente: "no me siento vivo de ninguna manera" y que la teoría gradualista equipara este reporte con aquél en el que otro sujeto expresa "un estado de nada, sin ningún estado de ánimo, como si estuviera muerto" (véase capítulo 1, sub-sección 1.4.2).

[59] Aunque me estoy centrando en el problema de la fenomenología, podríamos pensar que una ventaja de la teoría afectivista sobre la negacionista es que no tiene que recurrir a una explicación de cómo se puede acceder a la información de un estado sin recurrir a experiencias, pues hemos visto que los reportes de los pacientes sugieren que sus estados son fenoménicamente conscientes. Ahora, si de la interpretación afectivista se desprende que hay fenomenología parcial, entonces una explicación natural sobre el acceso podría ser que, al faltar fenomenología afectiva, el sujeto no tiene acceso a este aspecto afectivo, pero sí a otros aspectos. Por ejemplo, puede tener acceso a los objetos de la percepción, pero no acceso a los aspectos positivos o negativos que los acompañan. Siendo así, también hay acceso parcial.

No obstante, como sugiere la teoría afectivista respecto a este último reporte, si el sujeto está utilizando una metáfora o analogía para expresar la dificultad que tiene para experimentar estados afectivos, entonces es probable que el sujeto que reporta que no se siente vivo de ninguna manera (luego, se siente como muerto) también esté recurriendo a estos recursos lingüísticos para señalar la misma dificultad. Es decir, que no se siente vivo porque no puede experimentar emociones o estados de ánimo.

Esto resulta plausible si pensamos que solemos utilizar expresiones del tipo "estás muerto por dentro", por ejemplo, cuando vemos que alguien no siente empatía o no se conmueve por el sufrimiento ajeno[60]. Así, el paciente puede sentirse muerto si no es capaz de sentir emociones ante situaciones o personas que normalmente las causaban. Más aún, es posible que el sujeto no solo se sienta muerto por esta razón, sino también porque ya no disfruta las actividades cotidianas que solía disfrutar. Esto, a su vez, puede resultar en falta de interés o motivación para realizarlas.

La teoría afectivista, además de dar cuenta de la incapacidad de evocar emociones o respuestas emocionales, también puede dar cuenta del aspecto motivacional, pues ya vimos que las emociones también están compuestas por este aspecto. Por ejemplo, si te sientes feliz, no solo buscas opciones conductuales que mantengan el estado, sino que, por lo general, te sientes más motivado para realizar diferentes actividades, como salir con amigos o continuar con un trabajo. De manera similar, si estas actividades las encuentras placenteras o agradables (valencia positiva) esto te motiva para seguir fomentándolas. Ahora, si se pierden estos elementos afectivos (emociones o valencias asociadas con las actividades) también se puede perder el elemento motivacional asociado con ellos.

Ahora, ¿qué sucede con los reportes de consciencia e inconsciencia? Recordemos que uno de los pacientes reporta "simplemente me hundo en una especie de inconsciencia. Soy lo suficientemente consciente para saber que suceden cosas a mi alrededor, pero nada parece registrarse". Ante esto, la teoría negacionista parecía suponer que el paciente es suficientemente consciente para saber que suceden cosas a su alrededor porque tiene acceso

[60] Otras expresiones que utilizamos pueden ser "no tienes alma" o "estás vacío por dentro". En el Apéndice, el reporte 23 también podría señalar una deficiencia para experimentar estados afectivos, pues como puede apreciarse, el sujeto reporta que "es un cuerpo sin alma".

a la información de lo que sucede a su alrededor (o a la información que transmiten sus estados), pero se describe como inconsciente porque nada parece registrarse, suponiendo, a su vez, que esto significa que no hay nada que se siente para el sujeto ver lo que sucede a su alrededor, por ejemplo.

Más allá del problema general de suponer que el sujeto conoce las distinciones entre consciencia de acceso y consciencia fenoménica, señalaba que teóricos sofisticados como Billon y Kriegel saben muy bien que la fenomenología se define por cómo se siente para un sujeto estar en un estado o tener una experiencia, así que no tenían buenas razones para concluir la ausencia de fenomenología a partir del hecho de que nada parece registrarse. También mencioné que por "registro" podíamos entender procesamiento y/o recopilación de información y almacenamiento de ella. Por otra parte, podríamos preguntar cómo sabe el sujeto que suceden cosas a su alrededor. Una explicación natural es porque *percibe* lo que pasa en su entorno y esto es consistente con la tesis según la cual los sujetos tienen estados perceptivos (o están en dichos estados) que se sienten de cierta manera.

Siendo así, propongo que una mejor interpretación del reporte es la siguiente: cuando el sujeto dice que es suficientemente consciente para saber que suceden cosas a su alrededor se refiere a que puede percibir lo que sucede en su entorno y, tomando en cuenta lo anterior, podemos pensar que su percepción es fenoménicamente consciente[61], pero se describe como inconsciente o como "hundiéndose en una especie de inconsciencia" porque aquello que percibe no se registra o no le parece que se registra en el sentido que señalé anteriormente. Es decir, de alguna manera tiene problemas con la recopilación, almacenamiento y/o recuperación de la información.

Sin embargo, no debemos esperar que los sujetos utilicen específicamente estos términos, sino términos más cotidianos que, de cualquier manera, apuntan a estos procesos. Por ejemplo, pensemos que cuando terminamos de ver una película, suele suceder que no logramos recordar cierta información, ya sea de qué trató, de cómo se llamaba el personaje principal, de qué color era su cabello, etc. Si bien podríamos decir algo como "no tengo

[61] Esto no significa que debemos suponer que el sujeto sabe que su percepción es fenoménicamente consciente. Ya he establecido que es poco plausible que los sujetos ordinarios sepan qué significa esto. Solo a nivel teórico tenemos razones para pensar que lo es.

registro de lo que sucedió" (o incluso, "no tengo consciencia de lo que pasó"), es más común que expresemos "no me acuerdo de qué pasó". No obstante, en todos estos casos, estamos refiriéndonos a que no podemos recordar algo (o que no lo retenemos en la memoria) y este proceso involucra almacenamiento y/o recuperación de cierta información[62]. Ahora, la cuestión por resolver es cómo se relacionan los problemas de almacenamiento o recuperación de información con los aspectos afectivos que postula la teoría afectivista.

Regresemos al ejemplo de las películas; es más probable que te acuerdes de que trató la película si te causó un impacto emocional (si te entristeció la historia de amor o si te sorprendió el final), si amaste u odiaste a los protagonistas o si alguno de los personajes te pareció atractivo. En términos generales, es más fácil retener o recordar información si esta tiene una valencia o si causa un impacto emocional.

Más aún, autores como Brosch et al. (2013) han señalado que los estímulos emocionales o cargados afectivamente atraen *la atención* más fácilmente que los neutrales, provocando que te enfoques en ellos y prestes menos atención a información periférica. A su vez, esto resulta en una *mejor memoria* del estímulo con carga afectiva y en una mala memoria de la información periférica. "Un ejemplo de esto es 'el enfoque en el arma', la presencia de un arma en una escena resulta en una buena memoria de los detalles del arma y otros estímulos cercanos, pero en una mala memoria para otros detalles de la escena, como la cara del agresor" [2013, p. 4]. Luego, la atención a un estímulo con carga afectiva agudiza la memoria y/o facilita la retención y recuperación de los detalles de dicho estímulo[63].

A partir de lo anterior, podemos plantear que, cuando el sujeto despersonalizado reporta que es "suficientemente consciente", se refiere a que puede percibir lo que sucede a su alrededor, pero describe que se "hunde en una especie de inconsciencia", porque tiene dificultades para retener o recuperar información previamente vista, por ejemplo. En

[62] Por supuesto que la memoria es mucho más compleja y tiene diversas etapas. Una puede ser el almacenamiento de la información y otra la recuperación de ella. Sin embargo, cotidianamente no hacemos esas distinciones y cuando no nos acordamos de algo, no solemos decir que tenemos la información almacenada pero no podemos recuperarla. Normalmente nos podemos referir a ambas.

[63] De manera similar, Brady (2010, 2013) ha argumentado que parte del rol funcional de las emociones es enfocar la atención hacia aspectos relevantes de los objetos o de las situaciones. Retomando el ejemplo del enfoque en el arma, el miedo provocado ante la presencia del arma, atrae la atención hacia dicho objeto que se considera o se percibe como peligroso y, como sostienen Brosch y sus colegas, al enfocarse en este objeto, se presta menos atención a información periférica.

consecuencia, afirma que nada parece registrarse. A su vez, esta dificultad se debe a que ciertos estímulos han perdido la valencia o porque ciertas situaciones no causan un impacto emocional. Como sugieren Brosch y sus colegas, esto dificulta la atención y, por tanto, hay dificultades con la memoria o la recuperación de la información.

Si tomamos esta interpretación, evitamos el problema de suponer, no solo que el sujeto conoce las distinciones entre acceso y fenomenología, sino que también evitamos asimilar la fenomenología con el registro de información o, como veremos a continuación, con la comprensión de la información.

Recordemos que otro sujeto reportaba lo siguiente: "soy como un zombie, incapaz de asimilar ninguna información". Si es muy poco probable que el sujeto sepa qué es un zombie en el sentido que plantean los filósofos de la mente, ¿a qué se puede referir en el habla cotidiana cuando se describe como tal?

Para empezar, la mayoría de la gente tiene referencias de los zombies a partir de historias o películas de terror. En estas narrativas, los zombies son seres putrefactos con movimientos lentos, torpes y que comen gente, al parecer, sin ningún remordimiento. En relación con este último punto, los zombies en la cultura popular no parecen tener sentimientos (no solo se comen a cualquier persona sin remordimiento, sino que se comen a la gente que querían. Por ejemplo, familiares y amigos). Evidentemente, es poco probable que el sujeto se refiera a él mismo como zombie porque de pronto tiene el deseo de comer gente. Sin embargo, es probable que se compare o se describa como uno de ellos porque, de manera análoga a los zombies, es incapaz de tener emociones o sentimientos por alguien.

Si al lector no le convence esta interpretación, hay una alternativa. Perece ser muy común que la gente se describa como un zombie en ciertas circunstancias. Por ejemplo, como ya he señalado, cuando se siente lento y torpe, cuando se siente extremadamente cansado/agotado (no solo físicamente, sino también mentalmente) o cuando se siente somnoliento[64] y, aun sintiéndose así, tiene que realizar sus actividades cotidianas.

[64] De hecho, la lentitud y la torpeza pueden ser efectos del cansancio extremo y de la somnolencia. Por ejemplo, cuando estás agotado no tienes suficiente energía y esto puede disminuir la velocidad con la que te mueves o realizas tus actividades. Por otro lado, si estás somnoliento o no dormiste bien, entonces también eres más lento y actúas con torpeza.

La cuestión es, ¿cómo se relacionan estas sensaciones particulares con las dimensiones afectivas que se han propuesto? En el modelo circumplejo del afecto de Russell (1980, 2003), el autor identifica el cansancio extremo o el agotamiento con la fatiga y la sensación de somnolencia con el letargo y sostiene que estos son estados de activación/excitación baja o muy disminuida[65] (además, plantea que causan cierto grado de displacer o desagrado). Siendo así, si el sujeto se describe como un zombie debido a que se siente fatigado o aletargado y si estos estados se caracterizan principalmente por una activación disminuida, entonces la sensación de ser un zombie se puede explicar en términos de disminución de una dimensión afectiva.

Por otro lado, señalaba que la fatiga puede no ser únicamente física, sino también mental. Cuando hay fatiga o agotamiento mental, esto a menudo resulta en problemas de concentración; es difícil comprender lo que estamos leyendo, por ejemplo, o comprender lo que estamos diciendo o escuchando. Luego, la fatiga y/o la baja activación que experimenta el sujeto que se describe como un zombie, también puede explicar por qué el sujeto afirma que es incapaz de asimilar o comprender la información[66].

Por último, como ya señalé, si tomamos esta interpretación del reporte, evitamos asociar la fenomenología con la comprensión de la información. Por otra parte, si el sujeto se describe como un zombie porque se siente fatigado o aletargado, entonces hay algo que se siente para el sujeto sentirse así. Por tanto, podemos pensar que hay un estado fenoménicamente consciente.

Reportes que, en primera instancia, no apuntan a estados mentales: Hemos visto en los dos capítulos anteriores que, tanto la teoría gradualista como la teoría negacionista, plantean que son los estados de los pacientes los que no se sienten propios. No obstante, también

[65] Una cuestión por resolver es cuáles pueden ser los cambios fisiológicos asociados con estos estados. La falta de energía quizás se asocia con la disminución del flujo sanguíneo que puede resultar en la debilidad típica del cuerpo cuando nos sentimos agotados. La somnolencia quizás se asocia con la disminución de oxigenación y frecuencia cardiaca que puede resultar, en el caso de la oxigenación, en los bostezos constantes cuando estamos somnolientos.

[66] Otro reporte que puede explicarse en términos de fatiga y/o baja activación es aquel en el que el sujeto expresa “tengo la sensación de no tener ningún pensamiento” .También es común en el habla cotidiana utilizar expresiones del tipo “ya no puedo pensar más” o “tengo la mente en blanco” (expresiones similares que indican la impresión de que no hay pensamientos) cuando nos sentimos mentalmente agotados.

vimos que algunos reportes no parecen referir a ningún estado mental. Algunos reportes refieren a partes corporales y otros a movimientos automáticos.

Voy a comenzar por enunciar las dificultades que enfrentaban ambas teorías respecto a las partes corporales y cómo se pueden evitar o solucionar al adoptar una interpretación afectivista. Por otro lado, a partir de los problemas que enfrentan ambas teorías respecto a los movimientos automáticos, voy a señalar que la teoría afectivista también puede enfrentar dificultades para dar cuenta de estos reportes. Sin embargo, voy a ofrecer una hipótesis que, en principio, podría ayudar a solucionar dichas dificultades y que, a su vez, puede abrir nuevas líneas de investigación.

Respecto a las partes corporales, vimos que un sujeto reporta lo siguiente: "las partes de mi cuerpo se sienten como si no me pertenecieran" (el lector puede consultar el Apéndice o la sub-sección 1.4.2 del primer capítulo). En este caso, ambas teorías suponían que el estado que no se siente como si fuera propio es una sensación corporal. De esta manera, si la sensación corporal no se siente como si fuera suya (del paciente), como consecuencia, el sujeto sentiría que su parte corporal no le pertenece.

Sin embargo, de manera general señalé que, si esto fuera el caso, deberíamos suponer que la sensación corporal fundamenta la sensación de pertenencia del cuerpo. Es decir, si siento una sensación en el brazo, por ejemplo, entonces *necesariamente* siento que el brazo me pertenece. Si bien es cierto que típicamente dichas sensaciones están localizadas en nuestro propio cuerpo (o las sentimos en nuestro cuerpo), mencioné que los casos de somatoparafrenia mostraban que esto no siempre sucedía, pues algunos sujetos que padecen de esta condición reportan que *sienten* sensaciones en un cuerpo que no es suyo, sino de alguien más. En consecuencia, teníamos razones para pensar que una no fundamenta la otra o que no hay una relación necesaria entre ellas. Siendo así, aun si concediéramos que la subjetividad está atenuada o, incluso ausente de las sensaciones corporales, esto no necesariamente explica por qué el sujeto niega sus partes corporales[67].

[67] La teoría negacionista tiene una dificultad extra, ya que si las sensaciones corporales por definición *se sienten*, es difícil negar que sean fenoménicamente conscientes.

Ahora bien, si las sensaciones corporales no fundamentan la sensación de pertenencia del cuerpo, entonces cabe preguntar, ¿qué es aquello que fundamenta dicha sensación? Al responder a esta pregunta podremos ofrecer una interpretación alternativa de este reporte.

Frédérique de Vignemont (2018, 2023) ha ofrecido una teoría de la fenomenología de la pertenencia corporal en términos afectivos. Por motivos de extensión, es imposible abordar todos los postulados y los detalles de su teoría, así que sólo ofreceré algunos puntos clave de ella.

La autora propone lo que llama un *mapa corporal protector*. Este mapa no solo representa límites espaciales que determinan dónde sentimos las sensaciones corporales, sino que representa también qué límites corporales (o qué cuerpo en general) debe protegerse. Más aún, el cuerpo que debe protegerse tiene un *significado afectivo* especial para el sujeto y este tipo de afectividad la describe en términos narcisistas. Siguiendo a la autora, hay un sentimiento narcisista hacia el propio cuerpo que consiste en buscar su bienestar o lo mejor para él, como protegerlo de amenazas o de un daño potencial.

En términos generales, la tesis de la autora es que la pertenencia corporal tiene su base (o se fundamenta en) la representación afectiva del cuerpo o en el significado afectivo que tiene para uno mismo. En otras palabras, experimentar el cuerpo como propio o "ser consciente de que éste es mi cuerpo presupone que mi cuerpo tiene un significado afectivo especial para mí, y que mi cuerpo tenga tal significado es que sea el cuerpo que se debe proteger" [De Vignemont, 2023, p. 286]. Luego, si el cuerpo deja de tener dicho significado para un sujeto y/o si no se representa en el mapa corporal protector, entonces deja de sentirse como propio.

Ahora bien, ¿cómo se relaciona el significado afectivo del cuerpo para experimentarlo como propio con las dimensiones específicas que se han planteado a lo largo de este capítulo? Para la autora, la valencia de las sensaciones corporales es lo que ayuda a fijar los límites corporales que se deben proteger y/o es lo que ayuda a conferir un significado especial afectivo del cuerpo. Así, sostiene lo siguiente:

> [el mapa corporal protector] no representa simplemente límites espaciales, sino límites espaciales con una *valencia afectiva* (…) fijar estos límites implica determinar qué cuerpo tiene un valor específico. Es en este nivel afectivo donde pueden intervenir el dolor, la

> interocepción y otras sensaciones corporales con carga afectiva. Ni el tacto ni la visión tienen valencia afectiva intrínseca: no aportan información sobre el valor del cuerpo para el organismo. En cambio, el dolor confiere valencia a los límites corporales. Indica que estos límites son los que debemos cuidar y proteger [De Vignemont, 2018, pp. 186-187].

Para respaldar su hipótesis, la autora señala que los pacientes con somatoparafrenia no muestran aumento de la respuesta simpática cutánea (que, en términos generales, se mide por los cambios en la sudoración) cuando la parte corporal que niegan se ve amenazada. En cambio, cuando hay amenaza hacia extremidades que sí sienten propias, la respuesta simpática cutánea aumenta. Es decir, no hay activación fisiológica cuando hay amenaza hacia la extremidad "ajena" y no retiran la extremidad. Adicionalmente, estos pacientes reportan que, dado que no es su extremidad, no les importa lo que le suceda[68].

Ahora bien, si tomamos en cuenta la teoría de De Vignemont y, si los reportes de los pacientes con somatoparafrenia son similares al del paciente despersonalizado que siente como si sus partes corporales no le pertenecieran, entonces es posible[69] que el paciente no sienta su cuerpo como propio debido a que no tiene un significado afectivo para él. A su vez, porque cuando menos algunas de sus sensaciones corporales no están cargadas afectivamente o porque no están dotadas de valencia. Esto, como hemos visto con la interpretación de los dolores, no significa que no sientan la sensación en la parte corporal. La sienten, pero, de nuevo, al no tener una carga afectiva, no tiene un significado para él.

[68] Otro fenómeno donde la autora señala que sucede algo similar es en la insensibilidad congénita al dolor. Estos sujetos no solo muestran "ausencia total de incomodidad, muecas o retraimiento ante pinchazos prolongados, pellizcos en tejidos blandos y estímulos térmicos nocivos. Curiosamente, describen la sensación de que su cuerpo es un objeto externo o una especie de herramienta" [De Vignemont, 2018, p. 187].

[69] Por supuesto que esto es solo una posibilidad. De Vignemont señala, a partir de un estudio realizado por Dewe et al. (2016), que en la despersonalización tampoco hay respuesta afectiva cuando el cuerpo de los pacientes se ve amenazado. Sin embargo, debemos tomar la evidencia más cuidadosamente. En el estudio de Dewe y colegas se midió la respuesta cutánea en respuesta a amenazas hacia el cuerpo en sujetos *propensos a desarrollar experiencias disociativas* como las que se dan en la despersonalización y en la desrealización, no en sujetos que, de hecho, padecían de estas condiciones. Por otra parte, la respuesta cutánea solamente disminuyó, no desapareció por completo. Si bien estos hallazgos son relevantes y pueden proporcionar una mejor comprensión de la despersonalización y, específicamente, de la relación entre la afectividad y el propio cuerpo, necesitamos más y nueva evidencia. Por fortuna, la investigación sobre esta condición sigue en desarrollo.

Pasemos ahora a los reportes de movimientos automáticos: "notaba que mis manos y pies se movían, pero como si no me pertenecieran y se movieran automáticamente"[70] y "de pronto me pregunto: ¿realmente estoy aquí? ¿Realmente soy yo quien camina? Entonces hago enormes esfuerzos por aplicar mi consciencia a esta inconsciencia… para darme cuenta de que yo estoy haciendo los movimientos andantes. Entonces, en algún punto durante esta especie de crisis, antes de la absoluta certeza (de ser yo mismo, antes de la crisis) soy consciente por un lado de que soy inconsciente por otro lado".

Respecto al primer reporte, vimos que ambas teorías suponían que lo que explica la sensación de automatización es una intención para la acción. En el caso de la teoría gradualista, la intención no se siente claramente propia y, por tanto, al sujeto le parece como si sus manos y pies se movieran automáticamente. En el caso de la teoría negacionista, la intención no es subjetiva y, por tanto, es fenoménicamente inconsciente, pero accesible. No obstante, señalé que ninguna de las teorías explica u ofrece razones independientes de por qué debería ser una intención la que exhibe subjetividad atenuada o ausencia de subjetividad y no otro aspecto de la experiencia.

Por otro lado, en el segundo capítulo (véase sección 2.3) señalé que, si bien el segundo sujeto reporta que es consciente e inconsciente ("de pronto me pregunto: ¿realmente estoy aquí? ¿Realmente soy yo quien camina? Entonces hago enormes esfuerzos por aplicar *mi consciencia* a esta *inconsciencia*… para darme cuenta de que yo estoy haciendo los movimientos andantes"), el reporte sugiere que el paciente se refiere a que no le parece que él mismo realiza los movimientos andantes. Siendo así, mencioné que, entonces es similar al primer reporte. Luego, los negacionistas también podían suponer que, en este caso, el estado que no es subjetivo es la intención para la acción, pero dicha intención es accesiblemente consciente. Sin embargo, además de que no encontramos razones independientes para asumir lo anterior, señalé que tampoco encontrábamos nada en el reporte que sugiriera que el sujeto tiene una intención que no se siente suya y que lo incita a actuar de cualquier manera, así

[70] Si bien en este caso el sujeto también está reportando que siente que sus extremidades no le pertenecen, me voy a enfocar en la parte de los movimientos automáticos. Sin embargo, podemos pensar que la falta de pertenencia también se debe a que sus extremidades no tienen un significado afectivo para el sujeto y/o porque no se representan afectivamente y, *adicionalmente*, le parece que se mueven automáticamente. El mismo reporte parece sugerir dos dimensiones distintas.

como suele suceder con los reportes de control de los pacientes esquizofrénicos. Incluso, podemos aplicar el mismo punto para el primer reporte.

Por último, Sierra y David (2011) han argumentado que la sensación de automatización que reportan los pacientes despersonalizados es el resultado de una *falta de sentido de agencia*. Esto es, la sensación de que es uno mismo quién está causando o controlando una acción o movimiento. En la sub-sección 2.5.1 del segundo capítulo, vimos que esta interpretación tiene respaldo en la evidencia empírica, pues se ha encontrado que la CPI contribuye al sentido de agencia, de modo que, cuando hay una actividad anormalmente aumentada en esta área, el sentido de agencia desparece o disminuye[71].

Ahora bien, si la evidencia empírica respalda la interpretación de la ausencia del sentido de agencia, en la sub-sección 2.5.2 del segundo capítulo, vimos que los negacionistas podían suponer que la intención para la acción de alguna manera da lugar al sentido de agencia. Ya sea porque la intención para la acción no es más que el sentido de agencia mismo o, porque al no sentir la intención para la acción como propia, entonces el sujeto no siente que es el agente de sus movimientos. No obstante, no ofrecen una hipótesis clara que respalde ninguna de las dos opciones.

Si asumimos que la mejor interpretación para la sensación de automatización que reportan los pacientes es aquella que recurre a la falta de sentido de agencia, entonces la teoría afectivista también tendría problemas para dar cuenta de estos reportes, pues en principio, no parece haber una relación entre el sentido de agencia y las dimensiones afectivas que se han planteado a lo largo de este capítulo.

Sin embargo, voy a ofrecer una hipótesis que quizás podría relacionar el sentido de agencia con aspectos afectivos. La hipótesis es la siguiente: las experiencias afectivas o las dimensiones afectivas, como la valencia, pueden contribuir a la presencia o ausencia del sentido de agencia.

[71] Esta interpretación goza de amplia aceptación para dar cuenta de los movimientos automáticos, no solo en relación con la despersonalización, sino también en relación con la esquizofrenia. Como he señalado, estos pacientes suelen presentar delirios de control y, autores como Gallagher (2000), Farrer et al. (2004) y Frith et al. (2000), han argumentado que estos también se explican por la ausencia de sentido de agencia.

Consideremos las emociones. En la sección 3.3, vimos que Scarantino y De Sousa (2018) caracterizan las emociones como estados compuestos por diferentes aspectos, entre ellos, aspectos conductuales o motivacionales. En su ejemplo, cuando alguien siente miedo ante la presencia de un oso en su camino, hay una *tendencia a huir* del oso. Ahora, si te sientes triste por la ausencia de un ser querido, hay una *tendencia a llorar*. Si bien ambas emociones son negativas, ambas involucran diferentes tendencias de acción.

Lo mismo sucede con emociones positivas. Por ejemplo, si te sientes feliz porque tu equipo favorito ganó un partido, hay *una tendencia a saltar y gritar* y, si te sientes enamorado de alguien, hay una *tendencia a abrazarlo o besarlo*.

En todos estos casos hay tendencias típicas de cada emoción. No obstante, las emociones también pueden motivarnos a realizar actividades cotidianas. Por ejemplo, hemos visto que si te sientes feliz, por lo general te sientes más motivado para salir con amigos, para ir al trabajo o al colegio. Por otro lado, hay emociones o estados de ánimo que desmotivan[72]. Por ejemplo, cuando alguien se siente profundamente triste o deprimido, no se siente motivado para realizar ninguna actividad. Esta falta de motivación muchas veces se expresa como "falta de ganas" para hacer cualquier cosa.

Ahora bien, parece ser común que la gente que padece una profunda tristeza o un episodio depresivo y tiene que seguir con sus actividades cotidianas (como ir al trabajo o al colegio) sin ganas o motivación, suele expresar que tiene la impresión de ir en "piloto automático" o de realizar dichas actividades automáticamente[73].

Este parece ser un caso en el que una experiencia afectiva contribuye a una impresión de automatización. Esto es, la impresión de que no es uno mismo quien realiza sus actividades o quien controla sus acciones[74] (y quizás deberíamos añadir un "voluntariamente"). Si esto

[72] Es motivo de debate si todas las emociones involucran un elemento motivacional. Autores como Frijda (1986) y Scarantino (2014) ofrecen una respuesta positiva. Por ejemplo, si bien es cierto que la tristeza puede desmotivarnos para realizar ciertos cursos de acción cotidianos, involucran tendencias de acción de otro tipo. Como vimos, la tendencia a llorar.

[73] También suelen expresar que se sienten como si fueran robots. Los robots que conocemos hasta ahora, no tienen emociones y, en un sentido general, no parecen tener motivaciones. Son máquinas programadas para hacer lo que se tiene que hacer en automático.

[74] Si el lector alguna vez ha pasado por una situación similar, es probable que se sienta identificado con esta descripción.

sucede en casos cotidianos, es probable que también suceda en algunos casos de despersonalización, pues en el DSM-V se señala que las características que apoyan el diagnóstico, "incluyen diversos grados de ansiedad y depresión" [2014, p. 303]. Es decir, estos pacientes suelen presentar episodios depresivos.

Ahora, una cuestión que se podría plantear es si la falta de sentido de agencia (en la que contribuye una experiencia afectiva como la que acabo de señalar) es solamente una impresión de automatización o realmente una sensación. Es decir, si a los pacientes solo *les parece* que van en piloto automático o que realizan sus actividades automáticamente o, si literalmente, *sienten* que no son ellos quienes realizan sus acciones y/o actividades voluntariamente. Esta es una cuestión abierta que puede abrir nuevas líneas de investigación.

Consideremos ahora la valencia de otros estados, como los dolores. Respecto a la interpretación de estos estados, vimos que incorporan valencias negativas y que la ausencia de esta dimensión negativa provoca una ausencia motivacional, ya sea porque al no evaluar el dolor como malo o desagradable, no hay una motivación directa para retirarse de aquello que causa malestar o desagrado (Carruthers) o porque la valencia en sí misma es motivacional y, al faltar dicha valencia, no se siente la inclinación de deshacerse de la experiencia desagradable del dolor (Barlassina y Hayward).

En una línea similar, vimos que De Vignemont plantea un mapa corporal protector que representa límites corporales que deben protegerse y que el cuerpo tiene un significado afectivo especial para el sujeto que consiste en buscar lo mejor para él, como protegerlo de amenazas o de un daño potencial. Ahora, para proteger el cuerpo de amenazas o daños potenciales (como un dolor), se necesitan movimientos defensivos. Así, la autora sostiene que el sentido de pertenencia (que ya vimos que se fundamenta en el significado afectivo para el sujeto) "tiene una marca agentiva en el contexto de la auto-protección" [De Vignemont, 2023, p. 281]. Es decir, el mapa corporal "motiva a uno mismo a proteger el cuerpo que tiene tal significado" [Ibíd., p. 284].

Veíamos, además, que en el caso de la somatoparafrenia los sujetos no reaccionan cuando hay una amenaza hacia la extremidad que se siente ajena. No solo no hay activación fisiológica, sino que no retiran la extremidad. Esto es, no hay una motivación para proteger el cuerpo. Luego, se pierde la marca agentiva.

Ahora bien, en todos estos casos, si bien se pierde la motivación y/o la marca agentiva, los sujetos solamente no realizan las acciones correspondientes para protegerse del daño o retirarse para deshacerse del dolor. Es decir, no es que se muevan o actúen y tengan la impresión o la sensación de que no controlan sus movimientos o acciones.

Sin embargo, podemos considerar el reporte del paciente que afirma lo siguiente: "era doloroso y *sentía que mi brazo se retraía*, pero no era un dolor genuino (…)". En este caso, vimos que parecía haber ausencia de un componente motivacional, pues aunque el paciente sentía que su brazo se retraía, no indicaba que él mismo tuviera la disposición o la inclinación de retirarlo (voluntariamente). Si esto es el caso, podríamos pensar que el brazo se retraía *automáticamente* y/o que el paciente no controlaba voluntariamente el movimiento de su extremidad. Luego, este sería un caso en el que la ausencia de motivación (asociada con la ausencia de una valencia negativa) en ocasiones puede contribuir a la ausencia de sentido de agencia.

Una cuestión que podría plantearse a partir de este caso es si las inclinaciones o las tendencias de acción son siempre conscientes (ya sea fenoménicamente conscientes o accesiblemente conscientes) o si debemos separarlas de actos reflejos y, si es así, cómo se diferencian unas de otros. Si bien estas cuestiones están lejos del alcance de este trabajo, también podrían abrir nuevas líneas de investigación.

Por lo pronto, he ofrecido una hipótesis que puede relacionar el sentido de agencia con estados afectivos o con dimensiones afectivas específicas que, en principio, podrían ayudarnos a dar cuenta de los reportes de automatización de los pacientes despersonalizados. Como vimos, los estados afectivos (como la tristeza o la depresión) pueden contribuir a la falta de sentido de agencia, pues cuando uno está deprimido, no se siente motivado a realizar sus actividades cotidianas y, al tener que seguir realizándolas, se suele expresar que se tiene la impresión de realizarlas automáticamente o de ir en piloto automático. Por otro lado, vimos que la ausencia de motivación asociada con la ausencia de una valencia negativa en el caso del dolor puede contribuir a que los sujetos despersonalizados sientan que sus extremidades se retiren automáticamente.

En resumen, la teoría afectivista predice que, si hay déficits afectivos, los pacientes reportarán ausencia o disminución de emociones. Alternativamente, predice que, si diferentes

estados están compuestos por dimensiones afectivas específicas, como la valencia y la activación fisiológica, al haber ausencia o disminución de dichas dimensiones, los pacientes notarán que algo que normalmente estaba presente en su experiencia ya no está o parece haberse disminuido. En consecuencia, utilizarán expresiones no literales para tratar de explicar esta ausencia o disminución.

Como vimos, la mayoría de los reportes confirman las predicciones de la teoría. Si bien es cierto que, en el caso de la pertenencia corporal, los pacientes no parecen utilizar analogías o metáforas para describir la sensación de que sus partes corporales no son suyas, vimos que, de acuerdo con De Vignemont, dicha sensación es el resultado de un déficit afectivo. A saber, la incapacidad de representar su cuerpo afectivamente. Siendo así, esto no presupone mayor problema para la teoría.

A su vez, vimos que la interpretación de los reportes en términos de déficits afectivos nos permitió resolver o evitar muchos de los problemas que presentaban tanto la teoría gradualista como la teoría negacionista. Esta es una clara ventaja respecto a ambas teorías estructurales.

Ahora bien, vimos que si la mejor interpretación disponible para los reportes de automatización es aquella que recurre a la ausencia de sentido de agencia, entonces no solo las teorías estructurales presentan problemas para dar cuenta de estos reportes, sino también la teoría afectivista, pues, en principio, no parecía haber una relación entre el sentido de agencia y aspectos afectivos.

Sin embargo, ofrecí una hipótesis según la cual las experiencias afectivas o la dimensión afectiva de la valencia, podían contribuir a la presencia o ausencia del sentido de agencia. Si bien esta hipótesis está respaldada por ejemplos inicialmente plausibles, vimos que quedan cuestiones por resolver y que pueden explorarse en trabajos posteriores. Por lo pronto, esta es una hipótesis que puede ser prometedora.

En la siguiente sección, voy a mostrar que, a diferencia de las teorías estructurales, la evidencia empírica disponible converge con lo que postula la teoría afectivista. Lo cual representa también una ventaja ante estas teorías.

3.5. Evidencia empírica

Como vimos en la sección 1.6 del primer capítulo, los estudios de neuroimagen realizados a los pacientes despersonalizados muestran, principalmente, ausencia de actividad en la ínsula anterior y disminución de actividad en la amígdala e hipotálamo. Es ampliamente aceptado que estas áreas cerebrales desempeñan un papel fundamental en la generación de emociones y/o respuestas emocionales[75]. Por ejemplo, estudios realizados en monos han mostrado que los estímulos visuales deben llegar a la amígdala (o ser procesados en esta área) para que se produzca una respuesta emocional (Downer, 1961).

La amígdala no solo produce respuestas emocionales, sino que también se activa cuando hay estímulos cargados afectivamente[76], como imágenes emotivas tanto positivas como negativas (Shin et al., 1997 e Irwin et al., 1996), así como memorias emotivas (recordar un día especial con un ser amado, por ejemplo, o recordar la muerte de una persona cercana).

Por otra parte, Sierra y David (2011) han señalado que "estudios de resonancia magnética funcional han demostrado que la capacidad de experimentar sentimientos en respuesta a imágenes emocionales está directamente relacionada con la actividad de la ínsula anterior" [2011, p. 106] y, siguiendo a Morris (2002), sostienen que se ha observado activación de dicha área en relación con múltiples emociones, como la tristeza, la ira o el disgusto. De esta manera, concluyen lo siguiente: "esta amplia gama de correlatos de activación sugiere que el papel de la ínsula anterior en la generación de sentimientos es genérica en lugar de específica de una emoción en particular" [Ibíd.].

Por último, las respuestas fisiológicas de las emociones se han asociado con la actividad del hipotálamo. Hamann (2025) señala que en las ciencias afectivas y, especialmente en el ámbito neurocientífico, hay consenso respecto a las áreas encargadas de las "salidas emocionales fisiológicas". Entre ellas, el hipotálamo, que regula el sistema

[75] Es importante señalar que no solo estas áreas pueden estar involucradas en la generación de emociones. Otras áreas que entran dentro del llamado "sistema límbico" son el hipocampo, el tálamo, la circunvolución cingulada, entre otras. Por otra parte, puede haber controversia respecto a si son áreas emocionales específicas o si desempeñan otras funciones y/o si están involucradas en procesos no emocionales. (Véase Hamann, 2025).

[76] Recordemos que es posible distinguir entre *respuestas emocionales* y *percepción de estímulos cargados afectivamente* en tanto podemos percibir una imagen positiva o negativa, pero esto no implica *necesariamente* que experimentemos una emoción o una respuesta emocional cuando la vemos.

nervioso autónomo (por ejemplo, frecuencia cardiaca, sudoración, temperatura, respiración) y "el tronco encefálico, que tiene varias sub-áreas que controlan respuestas fisiológicas como la frecuencia cardiaca y la liberación de hormonas" [2025, p. 263].

Ahora bien, vimos que la nula actividad en la ínsula anterior se mostró en relación con imágenes aversivas o desagradables (valencia negativa) y que, adicionalmente, los pacientes reportaron que, si bien podían comprender el contenido de las imágenes, *no experimentaron ninguna emoción o respuesta emocional*. Lo cual es consistente con la teoría afectivista que predice ausencia de emociones.

También vimos que la reducción de actividad en la amígdala e hipotálamo se observó en relación con imágenes emotivas, como caras tristes y felices. Es decir, a diferencia de los controles, los pacientes despersonalizados no exhibieron una actividad normal en éstas áreas cuando veían imágenes que expresaban emociones positivas como negativas. En este caso, los resultados son consistentes con la predicción afectivista de disminución o atenuación de emociones, así como con la predicción de ausencia o disminución de aspectos afectivos. Por ejemplo, si la actividad de la amígdala aumenta (o si hay más actividad en esta área) cuando se perciben estímulos que se consideran valenciados o con carga afectiva, entonces una hipótesis podría ser que la incapacidad de percibir estímulos con valencias positivas o negativas está relacionada con una actividad gravemente disminuida en esta área.

Por otro lado, si el hipotálamo regula el sistema nervioso autónomo y/o las respuestas fisiológicas como la frecuencia cardiaca, temperatura, sudoración, cambios en respiración, etc., entonces la reducción de actividad en esta área puede resultar en la reducción de la activación fisiológica, lo cual puede contribuir a la atenuación de la intensidad de las emociones, así como en estados en los que hay una activación reducida, como cuando te sientes agotado tanto física como mentalmente o como cuando te sientes somnoliento[77].

Hay estudios que, de hecho, muestran disminución de dicha activación fisiológica. Por ejemplo, Sierra, Senior et al. (2002) han mostrado que los pacientes despersonalizados, a diferencia de los grupos de control, exhiben respuestas autonómicas reducidas (por ejemplo, en los niveles de conductancia de la piel) en respuesta a imágenes desagradables y,

[77] Y a nivel interpretativo, vimos que estos estados pueden estar relacionados con la sensación de ser un zombie que expresan los pacientes.

aunque señalan que no hay una diferencia significativa con los grupos de control, también hay cierta disminución en respuesta a imágenes agradables. Resultados similares muestran que, si bien estos pacientes no tienen mayor problema para calificar una imagen emotiva como agradable o desagradable, a diferencia de los grupos de control, sí muestran respuestas autonómicas reducidas (véase Sierra y David, 2011 y Sierra et al, 2002).

En suma, la evidencia empírica es consistente con la teoría afectivista que, en general, postula que las experiencias de los pacientes despersonalizados derivan de déficits afectivos.

Por último, en la sección 3.4, señalé que la teoría afectivista podía interpretar el reporte del paciente que se describe como *suficientemente consciente*, pero *hundiéndose en una especie de inconsciencia* de manera tal que podía percibir lo que sucedía a su alrededor, pero que nada parecía registrarse debido a que tenía dificultades para retener y/o recuperar información. Asimismo, siguiendo a Brosch et al. (2013), vimos que dichas dificultades derivaban de déficits afectivos, pues los autores plantean que *la percepción* de un estímulo emocional o con carga afectiva atrae *la atención* más fácilmente que la percepción de estímulos neutrales y que esto resulta en una *mejor memoria* (retención y/o recuperación) de los detalles de dicho estímulo.

Ahora, si el estímulo no está cargado afectivamente, entonces hay dificultades con la atención y esto resulta en una mala memoria del estímulo y/o de sus detalles. Esto es lo que podía suceder en el caso del paciente despersonalizado mencionado anteriormente.

Siendo así, parece que todos estos procesos (percepción, atención y memoria) se benefician y se relacionan en virtud de la información con carga afectiva. Vimos que un ejemplo de la relación de estos procesos que se benefician de los estímulos cargados afectivamente es el del enfoque del arma.

Adicionalmente, en relación con la percepción y la atención, los autores sostienen que diversas pruebas neuropsicológicas han mostrado que, cuando se percibe un estímulo emocional, no solo se facilita la atención al estímulo específico entre un número variado de distractores, sino que también se facilita la percepción de estímulos neutrales posteriores que aparecen en una misma ubicación.

> Estos efectos de orientación emocional se han demostrado mediante *la tarea de detección de puntos*, en la que los participantes responden a un estímulo específico que reemplaza una de dos señales presentadas simultáneamente; una con significado emocional (ej. una cara temerosa) y la otra neutral. Los resultados típicos muestran respuestas más rápidas a los estímulos que reemplazan la señal emocional que a la neutral. Las señales emocionales no solo pueden conducir a tiempos de detección más rápidos, sino que también aumentan directamente nuestra capacidad perceptiva al incrementar la sensibilidad al contraste para el objetivo subsiguiente [Brosch et al., 2013, p. 3].

Ahora bien, en la sub-sección 2.5.1 del segundo capítulo, vimos que los pacientes despersonalizados suelen ser incapaces de fijar la atención en un estímulo relevante (el que se les solicita indicar) en presencia de distractores. Aunado a esto, mencioné que Guralnik et al. (2000) han demostrado que estos pacientes también tienen *problemas de memoria* a corto plazo y postulan que estos problemas derivan de *problemas perceptivos* (disminución de sensibilidad perceptiva) y *atencionales*.

Luego, tenemos un contraste con lo que postulan Brosch y sus colegas (y con los estudios que mencionan) y los estudios y postulados de Guralnik et al. Esto se puede resumir de la siguiente manera: todo lo demás siendo igual, si el estímulo está cargado afectivamente, la percepción y la atención mejoran y esto resulta en una buena memoria del estímulo y sus detalles. En cambio, los estudios de Guralnik y colegas muestran problemas perceptivos y atencionales que pueden resultar en una mala memoria. Debido a que en estos estudios los estímulos no son emocionales, sino neutrales (recordemos que los estímulos son letras) esto podría dar respaldo a la tesis de Brosch et al.

Por supuesto que, para confirmarla, se tendrían que realizar estudios similares a los pacientes despersonalizados. Por ejemplo, algo similar al enfoque del arma o directamente someterlos a la tarea de detección de puntos. Hasta donde sabemos, esto no se ha hecho. Pero esto nos permite seguir motivando la investigación empírica de la despersonalización y el impacto de la percepción valenciada en otros procesos como la atención y la memoria.

3.6. Conclusión

De acuerdo a la teoría afectivista, las anomalías de las experiencias que reportan los pacientes despersonalizados son el resultado de *déficits afectivos* que pueden manifestarse de dos maneras: 1) como ausencia o disminución de emociones (o estados afectivos en general) o

2) como ausencia o disminución de dimensiones afectivas específicas que, generalmente, acompañan a muchos de nuestros estados mentales (no emocionales).

Al manifestarse como ausencia o disminución de emociones, esta teoría es similar a otras teorías afectivistas, pues, como vimos en la sección 3.2, estas plantean, de manera general, que en los pacientes despersonalizados hay falta de respuesta afectiva hacia personas, objetos y situaciones del mundo que normalmente la causaban y, señalé que esta falta de respuesta afectiva podía entenderse como falta o ausencia de emociones.

Sin embargo, en la sección 3.3, vimos que la teoría afectivista que planteo difiere de las anteriores en tanto no se restringe a una falta de respuesta afectiva/emocional (o disminución de respuesta afectiva/emocional) ante objetos, situaciones del mundo y personas que normalmente la causaban, sino que especifica dos dimensiones afectivas que no solo están presentes en los estados afectivos como las emociones, sino que también suelen acompañar a muchos de nuestros estados mentales, como pensamientos, percepciones y sensaciones corporales. Estas dimensiones corresponden a la *valencia* y *la activación fisiológica*.

Al identificar estas dimensiones afectivas vimos que la teoría afirma que, si otros estados incorporan dichas dimensiones, entonces al haber ausencia o disminución de valencia o activación fisiológica, los pacientes notarán que algo que normalmente estaba presente en sus experiencias ya no está o parece haberse disminuido y, en consecuencia, tratarán de explicar esta ausencia o disminución recurriendo a expresiones no literales.

A partir de lo que plantea la teoría afectivista, en la sección 3.4, ofrecí una interpretación alternativa de los reportes que consideran las teorías estructurales. A saber, la teoría gradualista y la teoría negacionista. Dicha interpretación no solo nos permitió resolver o evitar muchos de los problemas a los que se enfrentaban las teorías estructurales, sino que nos permitió interpretar otros aspectos de los reportes que ambas teorías pasaban por alto. Esto, sin duda, presenta una clara ventaja ante las teorías estructurales.

Si bien la gran mayoría de los reportes se pudieron interpretar en términos de déficits afectivos, ya sea como ausencia o disminución de emociones o como ausencia o disminución de valencia o activación fisiológica en otros estados (además de asociar estas dimensiones

con aspectos íntimamente relacionados, como los motivacionales), vimos que la teoría afectivista, al igual que las teorías estructurales, podía enfrentar problemas para dar cuenta de los reportes de automatización, ya que es ampliamente aceptado que la mejor interpretación para este tipo de reportes, es aquella que recurre a la ausencia de un *sentido de agencia*. Sin embargo, en principio, no parecía haber una relación entre el sentido de agencia y los aspectos afectivos que plantea la teoría.

Ante este problema, ofrecí una hipótesis según la cual las experiencias afectivas o la dimensión afectiva de la valencia pueden contribuir a la presencia o ausencia del sentido de agencia. Si bien esta es una hipótesis que puede ser prometedora debido a que está respaldada por ejemplos inicialmente plausibles, señalé que hay cuestiones por resolver que pueden desarrollarse en trabajos posteriores. Dicha hipótesis también puede representar una ventaja de la teoría afectivista ante las teorías estructurales, pues vimos que estas teorías afirman que lo que explica la sensación de automatización es una intención que no se siente propia (ya sea porque la subjetividad está atenuada o ausente). No obstante, vimos que no ofrecen razones independientes o una hipótesis que justifique por qué esto debería ser el caso.

Por último, en los dos capítulos anteriores, vimos que las teorías estructurales enfrentaban problemas con la evidencia empírica, pues esta no era consistente con lo que postulaban ambas teorías. En cambio, en la sección 3.5, mostré que otra ventaja de la teoría afectivista es que la evidencia empírica sí converge con lo que postula. Por un lado, vimos que los estudios de neuroimagen realizados a los pacientes despersonalizados han mostrado actividad anómala o nula actividad en áreas cerebrales que desempeñan un papel fundamental en la generación de emociones y/o respuestas emocionales, así como en áreas que se activan en presencia de estímulos cargados afectivamente (con valencia positiva o negativa) y en áreas que se relacionan con la activación fisiológica.

Por otro lado, vimos que estos pacientes exhiben problemas de atención y que, junto con la disminución de sensibilidad perceptiva, dichos problemas dan lugar a problemas con la memoria a corto plazo. Ahora, a partir de lo que afirman Brosch y sus colegas, vimos que la percepción de un estímulo cargado afectivamente, atrae la atención más fácilmente y que esto resulta en una mejor memoria del estímulo y sus detalles, de modo que, si el estímulo no está cargado afectivamente, entonces hay dificultades con la atención y esto resulta en una

mala memoria. Luego, podíamos concluir que, los problemas atencionales y de memoria, podían estar relacionados con déficits afectivos (la pérdida de valencia del estímulo, por ejemplo).

En suma, tanto las ventajas a nivel interpretativo como a nivel empírico que tiene la teoría afectivista ante las teorías estructurales, muestran que esta teoría ofrece una mejor interpretación de las anomalías de las experiencias que reportan los pacientes despersonalizados. Es decir, que las experiencias anómalas de estos pacientes derivan de déficits afectivos.

Conclusiones

Si las experiencias que reportan los pacientes despersonalizados son experiencias anómalas, entonces surgen preguntas filosóficas de suma importancia: ¿En qué consiste la anomalía de dichas experiencias? ¿Qué aspectos de las experiencias pueden fallar para que éstas sean anormales?

En este trabajo, identifiqué dos teorías que interpretan la anomalía de las experiencias en términos estructurales. Es decir, dichas teorías sostienen que la anomalía de las experiencias deriva de un aspecto general o estructural de ellas. A saber, del carácter subjetivo o subjetividad de la experiencia.

Por un lado, la teoría gradualista sostiene que, en los casos de despersonalización, la subjetividad de la experiencia está atenuada o disminuida en diferentes estados y, como consecuencia, las experiencias de estos sujetos se vuelven anómalas. Por otro lado, la teoría negacionista afirma que los estados de los pacientes son fenoménicamente inconscientes debido a que carecen de subjetividad. Sin embargo, al ser capaces de reportar dichos estados, sostiene que tienen acceso a la información que transmiten. En este caso, la anomalía consiste en que los sujetos son capaces de acceder a estados que no son subjetivos y que, por tanto, carecen de fenomenología propia.

Al interpretar las experiencias o estados de los pacientes en términos de atenuación de la subjetividad o de ausencia de subjetividad y, por tanto, de ausencia de fenomenología propia, ambas teorías buscan resguardar el principio subjetivo según el cual toda experiencia o estado fenoménicamente consciente es necesariamente subjetivo.

No obstante, argumenté que ambas teorías estructurales enfrentan dificultades de diferentes tipos y defendí una teoría no-estructural que evita o resuelve dichas dificultades. Ofrecí un argumento a la mejor explicación que dividí en dos pasos. En primer lugar, mostré que cada una de las teorías estructurales enfrenta tres dificultades diferentes para dar cuenta de la anomalía de las experiencias de estos pacientes. Si bien mi objetivo no fue refutar estas teorías, mostré que, en conjunto, tales dificultades eran lo suficientemente serias como para justificar la búsqueda de una teoría alternativa. En segundo lugar, esbocé una teoría afectivista que nos permitió evitar o resolver las dificultades identificadas en el primer paso

del argumento. De manera general, esta teoría sostiene que la anomalía de las experiencias no deriva de un aspecto estructural, sino de aspectos afectivos específicos.

En el primer capítulo, mostré que la teoría gradualista enfrenta dificultades interpretativas. Al señalar estas dificultades, argumenté que, para resolverlas, la teoría tendría que enriquecerse con hipótesis auxiliares. Pero, estas resultaron ser poco plausibles, poco específicas o, en otros casos, no se ofrecían razones independientes para adoptarlas.

Por otro lado, mostré que enfrenta dificultades teóricas, pues, pese a que hay diferentes concepciones de la subjetividad, suele aceptarse que ésta debe cumplir con dos características constitutivas para calificar como tal. Estas son pre-reflexividad y elusividad. Al recurrir a ciertos criterios de gradabilidad, argumenté que no tenemos razones para pensar que la subjetividad es una propiedad gradable, debido a que sus características constitutivas no satisfacen los criterios de gradabilidad.

Por último, mostré que también enfrenta problemas con la evidencia empírica disponible, pues esta no respalda lo que postula la teoría y/o es difícil de reconciliar con lo que afirma.

En el segundo capítulo, mostré que la teoría negacionista también enfrenta una serie de dificultades interpretativas que, para solucionarlas, también tendría que ser enriquecida con hipótesis auxiliares. No obstante, estas resultaron problemáticas por ser poco plausibles o porque van en contravía de lo que postula la teoría o de los mismos reportes.

Por otro lado, argumenté que esta teoría enfrenta dificultades explicativas, pues, en términos generales, no explica cómo se puede acceder a la información de un estado sin recurrir a experiencias o estados fenoménicamente conscientes. Luego, no ofrece una concepción del acceso que no explote la fenomenología del estado al que se accede.

Por último, mostré que también enfrenta dificultades con la evidencia empírica. Por una parte, la teoría sugiere que los pacientes no tienen dificultades con el acceso. Sin embargo, la evidencia neurológica muestra déficits atencionales que pueden dificultar o impedir el acceso. Por otra parte, al asumir que ciertas áreas cerebrales involucradas en la generación de ciertos estados individuales pueden contribuir a la fenomenología de dichos

estados, encontramos que la evidencia no converge con la tesis de ausencia de fenomenología.

En el tercer capítulo, esbocé una teoría afectivista que plantea que, la anomalía de las experiencias de los pacientes despersonalizados deriva de déficits afectivos específicos, ya sea que dichos déficits se presenten como ausencia o disminución de estados afectivos, como emociones o estados de ánimo, o como ausencia o disminución de dimensiones afectivas específicas, como la valencia y la activación fisiológica, que suelen acompañar a muchos de nuestros estados mentales (no emocionales).

La interpretación de los reportes en términos de déficits afectivos nos permitió evitar o resolver muchos de los problemas interpretativos a los que se enfrentaron las teorías estructurales. Entre ellos (aunque no exclusivamente), asumir supuestos poco plausibles, como suponer que los sujetos saben qué es un estado fenoménicamente consciente y uno fenoménicamente inconsciente y que, en consecuencia, cuando afirman que se sienten como si estuvieran muertos, se refieren a que se sienten fenoménicamente muertos (teoría gradualista). O, de manera similar, suponer que los sujetos conocen las distinciones entre consciencia de acceso y consciencia fenoménica y que, en consecuencia, se refieren a que son conscientes en el sentido de acceso e inconscientes en el sentido fenoménico y, por tal razón, se describen como zombies (teoría negacionista).

La teoría afectivista evita estos problemas al señalar que, en el habla cotidiana, es común que los sujetos utilicen analogías o metáforas para expresar cómo se sienten, pues, en muchas ocasiones, es difícil describirlo recurriendo al lenguaje literal. Así, afirma que los pacientes despersonalizados también suelen recurrir a estos recursos lingüísticos para expresar que hay un cambio en sus experiencias. Lo cual tiene respaldo en algunas descripciones de la despersonalización y en varios reportes de los pacientes. Por ejemplo, el sujeto que expresa que se siente como si estuviera muerto, también afirma que no tiene ningún estado de ánimo. En este caso, la interpretación afectivista es que el sujeto no se refiere literalmente a que está muerto, sino que es una expresión metafórica para indicar una deficiencia que consiste, precisamente, en la incapacidad de experimentar estados de índole afectivo.

De manera similar, la teoría afectivista interpreta el reporte del sujeto que se describe como zombie de manera tal que el sujeto está utilizando una analogía para señalar que se siente torpe y lento, así como retratan a los zombies en la cultura popular. Más aún, afirma que la lentitud y la torpeza son efectos del letargo y la fatiga que, en términos teóricos, son estados en los que hay disminución de activación fisiológica.

Ahora bien, la interpretación afectivista también nos permitió evitar los problemas empíricos a los que se enfrentan las teorías estructurales, pues, en términos generales y, a diferencia de estas teorías, la evidencia empírica sí converge con los postulados de la teoría. Esto es, las pruebas neurológicas han mostrado que hay daños en áreas cerebrales involucradas en la generación de emociones y/o respuestas emocionales (amígdala e ínsula anterior), en áreas que se activan en presencia de estímulos cargados afectivamente (amígdala) y en áreas que se relacionan con la activación fisiológica (hipotálamo).

Por último, parece que los déficits atencionales y de memoria que exhiben estos pacientes están directamente relacionados con déficits afectivos. Es decir, los déficits afectivos (como la pérdida de valencia de un estímulo), contribuyen a déficits atencionales y de memoria.

Así pues, las ventajas de la teoría afectivista ante las teorías estructurales nos muestran que esta teoría ofrece una mejor interpretación del carácter anómalo de las experiencias de los pacientes despersonalizados.

Cabe señalar que esta teoría también nos permitió ofrecer nuevas interpretaciones de los reportes que consideran otras teorías afectivistas. Por ejemplo, vimos que la teoría afectivista de Gerrans (2015) sostiene que la sensación de irrealidad que expresan algunos pacientes despersonalizados se debe a una falta de respuesta afectiva al mundo y que, de manera similar, Sierra y sus colegas (1998, 2002) sostienen que la sensación de irrealidad del mundo se debe a que la percepción de los objetos del mundo no está dotada (o impregnada) de sentimientos emocionales. Para ilustrar su punto, recurren a reportes de pacientes que padecen de hipoemocionalidad visual debido a que son similares a aquellos que expresan los pacientes despersonalizados.

No obstante, señalé que estos autores no explican qué significa que la percepción de los objetos del mundo no esté dotada de "sentimientos emocionales". Mencioné que una interpretación natural de esta expresión podría ser que tales sentimientos no los siente un sujeto al percibir los objetos del mundo. De ser así, podríamos pensar que el sujeto no siente emociones y/o que hay una falta de respuesta afectiva/emocional ante dichos objetos.

El problema es que, si esto es el caso, no es claro cómo se relaciona la ausencia de emociones y/o la falta de respuesta afectiva con los reportes de los pacientes con hipoemocionalidad visual y, por tanto, con los reportes de irrealidad de los pacientes despersonalizados (asumiendo que, de hecho, son similares), pues en sus reportes no hay nada que sugiera que no sienten emociones cuando ven ciertos objetos o escenas del mundo.

Ante esto, ofrecí una interpretación alternativa según la cual a los pacientes despersonalizados les parece que todo en la visión está muerto debido a que, de manera similar a los pacientes con hipoemocionalidad visual, los objetos les parecen extraños o irreales porque estos han perdido una cualidad esencial. A saber, una valencia positiva o negativa, sin que esto implique que, al no percibir la valencia de los objetos, ya no se produzca una respuesta emocional. Es decir, en estos casos, no es una cuestión de que ya no sientan emociones ante los objetos del mundo, sino que es una cuestión de que falta un aspecto que normalmente estaba incorporado en la percepción de los objetos y, al estar ausente, parecen como si fueran irreales o no auténticos.

Por otro lado, Sierra y sus colegas no ofrecen una interpretación de los reportes que refieren a la sensación de ser un zombie. Si bien este tipo de reporte se puede interpretar de manera tal que, de manera análoga a los zombies, los sujetos no sienten emociones hacia ninguna persona, ofrecí una interpretación novedosa según la cual la sensación de ser un zombie se explica por una activación fisiológica disminuida, como en los casos en los que nos sentimos aletargados o fatigados física y mentalmente y, en consecuencia, nos podemos describir como si fuéramos zombies.

De manera similar, los reportes que expresan la sensación de no tener pensamientos pueden explicarse en términos de fatiga mental (producto de una baja activación fisiológica), pues es común utilizar expresiones análogas como "ya no puedo pensar más" o "tengo la mente en blanco" cuando nos sentimos mentalmente agotados.

Por último, la interpretación de Sierra y sus colegas de los reportes de automatización no recurre a ningún déficit afectivo. En cambio, recurre a problemas con el sentido de agencia. Es decir, sostienen que, en estos casos, hay una ausencia de dicho sentido. Vimos que esta interpretación goza de amplia aceptación, no solo en relación con la despersonalización, sino también en relación con los delirios de control que expresan algunos pacientes esquizofrénicos. Además, esta interpretación está respaldada empíricamente.

Desafortunadamente, si esto es el caso, señalé que la teoría afectivista que propongo parece enfrentar problemas para dar cuenta de estos reportes, pues no parece haber una relación entre el sentido de agencia y aspectos afectivos.

No obstante, ofrecí una hipótesis según la cual las experiencias afectivas o la dimensión afectiva de la valencia pueden contribuir a la presencia o ausencia de la sensación de sentido de agencia. Dicha hipótesis está respaldada por ejemplos inicialmente plausibles. Por ejemplo, parece ser común que, cuando la gente se siente deprimida y, por tanto, sin ganas o motivación para realizar actividades cotidianas, suele expresar que tiene la impresión de ir en piloto automático. En este caso, si bien es plausible suponer que una experiencia afectiva contribuye a una falta de sentido de agencia, se necesita desarrollar a detalle el papel específico de las emociones (y sus aspectos constitutivos, como los motivacionales) en la experiencia de agencia. Otra cuestión por resolver es si los pacientes despersonalizados realmente sienten que sus movimientos o acciones son automáticos o si sólo es una impresión de automatización.

Por otro lado, señalé que el sujeto que expresa que sentía que su brazo se retraía (al sentir dolor), no parecía tener una disposición o inclinación a retirarlo voluntariamente. Luego, podíamos pensar que el brazo se retiraba automáticamente y, de ser así, señalé que la ausencia del aspecto negativo del dolor (valencia negativa), resultaba en una ausencia motivacional para retirar el brazo voluntariamente. En consecuencia, la falta de una dimensión afectiva en ocasiones podía contribuir a una ausencia de sentido de agencia. No obstante, a partir de este caso, podíamos preguntarnos si las inclinaciones o las tendencias de acción (que corresponden al aspecto motivacional) son necesariamente conscientes, tanto en el sentido fenoménico como en el sentido de acceso o si son distintas de actos reflejos y, de ser así, cómo podríamos distinguirlas.

Si bien todas estas son cuestiones abiertas y/o que necesitan desarrollarse, pueden abrir nuevas líneas de investigación.

Apéndice de reportes

1) “Las partes de mi cuerpo se sienten como si no me *pertenecieran*” [Sierra y Berrios, 2000, p. 160]

2) “Parece que estoy caminando en un mundo que reconozco pero que no siento” [Sierra, 2009, p. 32]

3) “Besar a mi esposo es como besar una mesa, señor. La misma cosa… ni la menor emoción. Nada en la tierra puede emocionarme. Ni mi esposo ni mi hijo…mi corazón no late. No puedo sentir nada” [Dugas y Moutier, 1911, p. 109]

4) “Puedo sentirme entumecido de sentimientos, casi vacío por dentro. Odio el hecho de que no puedo sentir las cosas como solía hacerlo” [Sierra, 2009, p. 27]

5) “Tengo la sensación de no tener ningún pensamiento” [Sierra y Berrios, 2000, p. 161]

6) “No me siento vivo de ninguna manera” [Sierra, 2009, p. 8]

7) “Un estado de nada, sin ningún estado de ánimo, como si estuviera muerto” [Simeon y Abugel, 2006, p. 30]

8) “ella simplemente siente como si no existiera” [Simeon y Abugel, 2006, p. 8]

9) “A menudo tengo que hacerlo… entrar en una tienda para hablar, para pedir algo, para obtener una prueba de que soy realmente yo mismo” [Séglas y Meige, 1895, p. 131]

10) “Debo ser alguien, soy alguien, todos los demás se sienten alguien, pero yo no soy yo mismo. De repente sentí que estaba en la mitad, hay dos de mí” [Shorvon, 1946, p. 782]

11) “Me siento tan desapegado de mis pensamientos que parecen tener una ‘vida’ propia” [Sierra y Berrios, 2000, p. 163]

12) “Notaba que mis manos y pies se movían, pero como si no me *pertenecieran* y se movieran automáticamente” [Sierra, 2009, p. 29]

13) “Cuando una parte de mi cuerpo duele, me siento tan despegado del dolor que se siente como si fuera el dolor de alguien más” [Sierra y Berrios, 2000, p. 163]

14) “Era doloroso y sentí que mi brazo se retraía, pero no era un dolor genuino, era un dolor que no llegaba al alma… es un dolor, si quieres, pero la superficie de mi piel está tres kilómetros lejos de mi cerebro y no sé si estoy sufriendo” [Janet, 1928, p. 65]

15) “Cada uno de mis sentidos, cada parte de mi propio yo es como si estuviera separado de mí y ya no pudiera proporcionarme ninguna sensación… mis ojos ven y mi espíritu percibe, pero la sensación de lo que veo está completamente ausente” [Sierra, 2009, p. 8]

16) “Literalmente no había más experiencia de ‘mi’ en lo absoluto. La experiencia de la identidad personal se apagó y no volvería a aparecer jamás… el cuerpo, la mente, el habla, los pensamientos y las emociones, estaban todos vacíos; *no tenían pertenencia*, ninguna persona detrás de ellos” [Simeon y Abugel, 2006, pp. 143-44]

17) “De pronto me pregunto: ¿realmente estoy aquí? ¿realmente soy yo quien camina? Entonces hago enormes esfuerzos por aplicar mi consciencia a esta inconsciencia… para darme cuenta de que yo estoy haciendo los movimientos andantes. Entonces, en algún punto durante esta especie de crisis, antes de la absoluta certeza (de ser yo mismo, antes de la crisis) soy consciente por un lado de que soy inconsciente por otro lado” [Séglas y Meige, 1895, p. 147]

18) “Todo en la visión está muerto” [Mayer-Gross, 1935, p. 111]

19) “Puedo ver, oír y oler, pero es como si no viera, oyera u oliera” [Dugas y Moutier, 1911, p. 10]

20) “La parte emocional de mi cerebro está muerta… todas mis emociones están aplanadas” [Shorvon, 1946, p. 784]

21) “Simplemente me hundo en una especie de inconsciencia. Soy los suficientemente consciente para saber que suceden cosas a mi alrededor, pero nada parece registrarse” [Shorvon, 1946, p. 784]

22) “Soy como un zombie, incapaz de asimilar ninguna información” [Sierra, 2009, p. 51]

23) “Es un estado en el que no sientes nada, en el que no piensas, en el que no quieres decir lo que haces o piensas… estoy en el vacío, soy un cuerpo sin alma…veo sin ver, soy un hombre ciego que ve” [Janet, 1928, pp. 51-52]

24) “Es la sensibilidad mental la que falta, no soy yo quien siente. No tengo interés en lo que parezco estar sintiendo. Es otra persona la que siente mecánicamente” [Janet, 1908, p. 515]

25) “Puedo sentarme mirando mi pie y mi mano y no sentir que son *míos*. Esto puede suceder cuando estoy escribiendo, mi mano solo está escribiendo pero no se lo digo. Casi se siente como si hubiera muerto, pero nadie me lo ha dicho. Entonces me quedo viviendo en un caparazón que ya no reconozco” [Sierra, 2009, p. 27]

26) “Era *como si no fuera yo* el que caminaba, como si no fuera yo el que hablaba, como si no fuera yo el que vivía (…) puedo mirarme, de alguna manera me molesta mi cuerpo, como si no fuera yo, como si viviera a lado de mi cuerpo, de lado de mí mismo si quieres. No sé cómo explicarlo” [Janet y Raymond, 1898, p. 70]

27) “Solo siento ira por fuera, por sus reacciones fisiológicas” [Dugas y Moutier, 1911, p. 121]

28) “Tengo miedo de mis pensamientos, de mi cerebro, etc. Pero fundamentalmente no tengo miedo en absoluto. Si tan solo pudiera tener un miedo genuino” [Dugas y Moutier, 1911, p. 121]

Referencias

Albahari, M. (2006). *Analytical Buddhism: The two-tiered illusion of Self*. New York: Palgrave Macmillan.

Armstrong, D.M. (1962). *Bodily sensations*. London: Routledge and Kegan Paul.

Asociación Americana de Psiquiatría (2014). *Manual diagnóstico y estadístico de los trastornos mentales* (DSM-5), 5a Ed. Arlington, VA, Asociación Americana de Psiquiatría.

Aydede, M. (2019). Pain. *The Stanford Encyclopedia of Philosophy.* https://plato.stanford.edu/entries/pain/

Barlassina, L. and Hayward, M.K. (2019). More of me! Less of me! Reflexive imperativism about affective phenomenal character. *Mind*, 128 (512): 1013-1044.

Barlassina, L. (2020). Beyond good and bad: Reflexive imperativism, not evaluativism, explains valence. *Thought: A Journal of Philosophy*, 9: 274-284.

Baars, B.J. (1988). *A Cognitive Theory of Consciousness.* Cambridge University Press.

Barret, L.F. and Lida, T. (2025). "Constructionist theories of emotion in psychology and neuroscience". in Scarantino, A (ed.). *Emotion Theory: The routledge comprehensive guide. Vol. I: History, contemporary theories and key elements.* New York: Routledge.

Brady, M. (2010). Virtue, Emotion, and attention. *Metaphilosophy*, 41 (1): 115-131.

Brady, M. (2013). *Emotional Insight: The epistemic role of emotional experience*. Oxford: Oxford University Press.

Bayne, T. (2010). *The Unity of Consciousness.* New York: Oxford University Press.

Billon, A. (2013). Does consciousness entail subjectivity? The puzzle of thought insertion. *Philosophical Psychology*, 26 (2): 1-24.

Billon, A. (2016). Making sense of the Cotard syndrome: insights from the study of depersonalisation. *Mind and Language*, 31 (3):1-53.

Billon, A. (2023). “What Is It Like to Lack Mineness? Depersonalization as a probe for the scope, nature and role of mineness.” in García-Carpintero, M and Guillot, M (eds.) *Self-experience*: *Essays on Inner Awareness.* Oxford University Press.

Billon, A. and Kriegel, U. (2015). “Jaspers´ Dilemma: The Psychopathological Challenge to Subjectivity Theories of Consciousness” in Gennaro, R (ed.) *Disturbed Consciousness*: *New Essays on Psychopathology and Theories of Consciousness.* MIT Press.

Block, N. (1995/2007). “On a Confusion about a Function of Consciousness” in Block, N. (ed.) *Consciousness, Function and Representation: Collected papers,* Vol. 1. MIT Press Bradford Books.

Block, N. (2003). “Mental Paint” in Hahn, M. & Ramberg, B. (eds.) *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. MIT Press.

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain sciences,* 30 (5-6): 481-548.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in cognitive sciences,* 15 (12):567-575.

Bockner, S. (1949). The depersonalization syndrome: report of a case. *Journal of Mental Science,* 95: 968-971.

Brosch, T., Scherer, K.R., Grandjean, D., Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss Medical Weekly,* Vol. 143: 1-10.

Campbell, J. (1999). Schizophrenia, the Space of Reasons and Thinking as a Motor Process. *The monist* 82: 609-625.

Carruthers, P. (2018). Valence and value. *Philosophy and Phenomenological Research,* 97 (3): 658-680.

Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. Oxford University Press (USA).

Chudnoff, E. (2015). *Cognitive Phenomenology*. London: Routledge.

Colombetti, G. and Kuppens, P. (2025). “How should we understand valence, arousal, and their relation?” in Scarantino, A (ed.). *Emotion Theory: The routledge comprehensive guide. Vol. I: History, contemporary theories and key elements.* New York: Routledge.

Critchley, H.D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *Journal of Comparative Neurology,* 493: 154-166.

Dehaene, S., Kerszberg, M. and Changeux, J.P. (1998). A Neuronal Model of a Global Workspace in Effortful Cognitive Tasks. *Proceedings of the National Academy of Sciences of the United States of America*, 95 (24): 14529–14534.

Dehaene, S. and Naccache, L. (2001). Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework. *Cognition*, 79 (1–2): 1–37.

Dehaene, S., Changeux, J.P., Naccache, L., Sackur, J. and Sergent, C. (2006). Conscious, Preconscious, and Subliminal Processing: A Testable Taxonomy. *Trends in Cognitive Sciences*, 10 (5): 204–11.

Dehaene, S. (2009). “Neural Global Workspace”. in Bayne, T. Cleeremans, A. and Wilken, P. (eds.) *The Oxford Companion to Consciousness*. Oxford University Press.

De Vignemont, F. (2018). *Mind the body: An exploration of bodily self-awareness.* Oxford: Oxford University Press.

De Vignemont, F. (2023). “The phenomenology of bodily ownership”. in García-Carpintero, M and Guillot, M (eds.) *Self-experience*: *Essays on Inner Awareness.* Oxford University Press.

Dewe, H., Watson, D.G. and Braithwaite, J.J. (2016). Uncomfortably numb: new evidence for suppressed emotion in response to body-threats in those predisposed to sub-clinical dissociative experiences. *Cognitive Neuropsychiatry*, 28: 1-25.

Downer, J.D.C. (1961). Changes in visual gnostic functions and emotional behavior following unilateral temporal lobe damage in “split brain” monkey. *Nature*, 191: 50-51.

Dugas, L. and Moutier, F. (1911). *La dépersonnalisation*. Paris: F. Alcan.

Duncan, M. (2018). Subjectivity as Self-Acquaintance. *Journal of Consciousness Studies*, 25, No. 3–4, 2018, pp. 88–111.

Farah, M. (1994). "Visual perception and visual awareness after brain damage: A tutorial overview" in Umilta, C. and Moscovitch, M. (eds.) *Consciousness and unconscious information processing: attention and performance 15*. MIT Press. pp. 203-236.

Farrer, C. and Frith, C. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage,* 15, 596-603.

Farrer, C., Franck, N., Georgieff, N., Frith, C., Decety, J. and Jeannerod, M. (2003). Modulating the experience of agency: A positron emission tomography study. *Neuroimage*, 18, 324-333.

Farrer, C., Franck, N., Frith, C., Decety, J., Georgieff, N., d´Amato, T. et al. (2004). Neural correlates of action attribution in schizophrenia. *Psychiatry Research,* 131, 31-44.

Frijda, N.H. (1986). *The emotions*. Cambridge: Cambridge University Press.

Frith, C., Blakemore, S. J. and Wolpert, D. M. (2000). *Abnormalities in the awareness and control of action. Philosophical Transactions of the Royal Society of London Series B Biological Sciences*, 355, 1771-1788.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive sciencie. *Trends in Cognitive Sciences* 4 (1): 14-21.

Gallagher, S. (2004). Neurocognitive models of schizophrenia: A neurophenomelogical critique. *Psychopathology*, 37 (1): 8-19.

Gennaro, R. (2015). "Psychopathologies and Theories of Consciousness: An overview" in Gennaro, R (ed.) *Disturbed Consciousness*: *New Essays on Psychopathology and Theories of Consciousness.* MIT Press.

Gennaro, R. (1996). *Consciousness and self-consciousness*. Amsterdam: John Benjamins.

Gerrans, P. (2015). "All the self we need". in Metzinger, T and Windt, J.M (eds.) *Open Mind.* Mind Group.

Graham, G. (2010). *The Disordered Mind: An Introduction to Philosophy of Mind and Mental Illness*. London: Routledge.

Guillot, M. (2016). I me mine: on a confusion concerning subjective character. *Review of philosophy and psychology* 8 (1), 23-53.

Guralnik, O., Schmeidler, J. and Simeon, D. (2000). Feeling unreal: Cognitive Processes in depersonalization. *Am J Psychiatry,* 157 (1):103-109.

Hamann, S. (2025). "An overview of contemporary theories of emotions in neuroscience". in Scarantino, A (ed.). *Emotion Theory: The routledge comprehensive guide. Vol. I: History, contemporary theories and key elements.* New York: Routledge.

Hawthorne, J. and Logins, A. (2021). Graded Epistemic Justification. *Philosophical Studies* vol. 178, 1845-1858.

Hume, D. (1740/1888). *A Treatise of Human Nature,* L.A. Selby-Bigge (ed.) Oxford University Press.

Irwin, W., Davidson, R.J., Lowe, M.J., Mock, B.J., Sorenson, J.A., Turski, P.A. (1996). Human amygdala activation detected with echoplanar functional magnetic resonance imaging. *Neuroreport,* 11: 1765-1769.

Janet, P. and Raymond, F. (1898). *Névroses et idées fixes, Vol. 2.* Paris: Felix Alcan.

Janet, P. (1908). Le sentiment de dépersonnalisation. *Journal de Psychologie Normale et Pathologique,* 5: 514-516.

Janet, P. (1928). *De l'angoisse à l'extase. Vol. 2, Le sentiments fundamentaux.* Paris: Felix Alcan.

Kennedy, C. (1999). *Projecting the adjective: The syntax and semantics of gradability and comparison*. New York: Garland.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy,* 30 (1) 1-45.

Kriegel, U. (2004). Consciousness and self-consciousness. *The monist,* 185-209.

Kriegel, U. (2009). *Subjective Consciousness: A self-Representational Theory*. New York: Oxford University Press.

Lee, A. (2022). Degrees of Consciousness. *Noûs* 1-23.

Lebrecht, S., Bar, M., Barret, L.F., Tarr, M.J. (2012). Micro-valences: perceiving affective valence in everyday objects. *Frontiers in Psychology*, 3: 107.

Lemche, E., Surguladze, S. A., Giampietro, V. P., Anilkumar, A., Brammer, M. J., Sierra, M. et al. (2007). Limbic and prefrontal responses to facial emotion expressions in depersonalization. *Neuroreport*, 18, 473-477.

Lemche, E., Anilkumar, A., Giampietro, V. P., Brammer, M. J., Surguladze, S. A., Lawrence, N. S. et al. (2008). Cerebral and autonomic responses to emotional facial expressions in depersonalization disorder. *British Journal of Psychiatry*, 193, 222-228.

Levine, J. (2001). *Purple haze: The puzzle of consciousness*. New York: Oxford University Press.

Mack, A. and Rock, I. (1998). *Inattentional blindness.* Cambridge, MA: MIT Press.

Mayer-Gross, W. (1935). On depersonalization. *British Journal of Medical Psychology* 15: 103-126.

Morales, J. (2023). Mental strength: A theory of experience intensity. *Philosophical Perspectives*, 37: 1-21.

Morris, J.S. (2002). How do you feel? *Trends in Cognitive Sciences*, 6: 317-319.

Morrison, J. (2016). Perceptual confidence. *Analytical Philosophy*, Vol. 57 N° 1: 15-48.

Phillips, I. (2018). Unconscious perception reconsidered. *Analytic philosophy*, 59 (4): 471-514.

Pitcher, G. (1970). Pain perception. *The Philosophical Review*, 79 (3): 368-393.

Rosenthal, D. (1993). State Consciousness and Transitive Consciousness. *Consciousness and Cognition* 2: 355-363.

Rosenthal, D. (1997). “A Theory of Consciousness” in Block, N., Flanagan, O. and Güzeldere (eds.) *The Nature of Consciousness: Philosophical Debates. Cambridge,* MA: MIT Press.

Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, Vol. 39, N° 6: 1161-1178.

Russell, J.A. (2003). Core affect and the psychological construction of emotion. *Psychological Review,* Vol. 110, N° 1: 145-172.

Satpute, A.B., Kang, J., Bickart, K.C., Yardley, H., Wager, T.D., Barret, L.F. (2015). Involvement of sensory regions in affective experience: a meta-analysis. *Frontiers in Psychology*, 6.

Scarantino, A. (2014). “The motivational theory of emotions”. in Jacobson, D and D´Arms, J (eds.). *Moral psychology, and human agency.* New York: Cambridge University Press.

Scarantino, A. and De Sousa, R. (2018). Emotion. *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/emotion/

Sebastián, M. (2012). Experiential Awareness: Do you Prefer It to Me? *Philosophical Topics*, Vol 40. N° 2: 155-177.

Sebastián, M. (2014). Dreams: An empirical way to settle the discussion between cognitive and no-cognitive theories of consciousness. *Synthese*, 191 (2):263-285.

Sebastián, M. (2020). “Perspectival self-consciousness (as core De Se) and Ego-Dissolution: An analysis of (some) altered states of consciousness “in Metzinger, T. and Millière, R. (eds.) PhiMiSci Special Issue: *Radical Disruptions of Self-Consciousness*.

Sebastián, M. (2022). *El problema de la consciencia: Una introducción crítica a la discusión filosófica actual*. España: Cátedra.

Séglas, J. and Meige, H. (1895). *Leçons cliniques sur les maladies mentales et nerveuses (Salpêtrière 1887-1894)*. Paris: Asselin & Houzeau.

Shin, L.M., Kosslyn, S.M., McNally, R.J., Alpert, N.M., Thompson, W.L., Rauch, S.L., et al. (1997). Visual imagery and perception in postraumatic stress disorder. *Archives of General Psychiatry,* 54: 233-241.

Shorvon, H.J. (1946). The depersonalization syndrome. *Proceedings of the Royal Society of Medicine,* 39: 779-792.

Shuman, V., Sander, D. and Scherer, K.R. (2013). Levels of valence. *Frontiers in Psychology,* 4: 261.

Sierra, M. (2009). *Depersonalization: A new look at a neglected syndrome.* Cambridge University Press.

Sierra, M. and Berrios, G. (1998). Depersonalization: Neurobiological Perspectives. *Biol. Psychiatry*, Nov 1; 44(9):898-908.

Sierra, M. & Berrios, G. (2000). The Cambridge depersonalization scale: a new instrument for the measurement of depersonalization. *Psychiatry research,* 93 (2), 153-164.

Sierra, M., Lopera, F., Lambert, M. V., Phillips, M. L. and David, A. S. (2002). Separating depersonalization and derealisation: the relevance of the "lesion method". *J Neurol Neurosurg Psychiatry,* 72: 530-532.

Sierra, M., Senior, C., Dalton, J., McDonough, M., Bond, A., Phillips, M.L., et al. (2002). Autonomic response in DPD. *Archives of General Psychiatry,* 59: 833-838.

Sierra, M. and David, A. S. (2011). Depersonalization: A selective impairment of self-awareness. *Consciousness and Cognition* 20: 99-108.

Simeon, D., Guralnik, O., Hazlett, E. A., Spiegel-Cohen, J., Hollander, E. and Buchsbaum, M. S. (2000). Feeling unreal: A PET study of depersonalization disorder. *American Journal of Psychiatry,* 157, 1782-1788.

Simeon, D. and Abugel, J. (2006). *Feeling Unreal: Depersonalization Disorder and the Loss of the Self.* Oxford: Oxford University Press.

Strawson, G. (2011). "Cognitive Phenomenology: real life" in Bayne, T. & Montague, M. (eds.) *Cognitive Phenomenology*. Oxford University Press.

Tye, M. (2006). “Another look at representationalism about pain”. in Aydede, M (ed.) *Pain: New essays on its nature and the methodology of its study.* Cambridge: MIT Press.

Weiskrantz, L. (1986). *Blindsight: A case study and implications*, 1a ed., Oxford University Press (USA).

Williford, K. (2015). “Representationalisms, subjective character and self-acquaintance.” in Metzinger, T. and Windt, J.M (eds.) *Open mind*. Mind group.

Wu, W. (2014). *Attention*. New York: Routledge.

Wu, W. and Morales, J. (2024). The Neuroscience of Consciousness. *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/consciousness-neuroscience/

Zahavi, D. (2005). *Subjectivity and selfhood: Investigating the first-person perspective*. Cambridge: MIT Press.

Zahavi, D. (2011). “Unity of consciousness and the problem of self” in Gallagher, S. (ed). *The Oxford Handbook of the Self.* Oxford University Press.

Zahavi, D & Kriegel, U. (2015). “For-me-ness: what it is and what it is not” in A.E. W.H. D. Dahlstrom (ed.), *Philosophy of mind and phenomenology: conceptual and empirical Approaches.* London: Routledge.

Zajonc, R.B. (1980). Feeling and thinking: preferences need no inferences. *American Psychologist*, Vol. 35, N°2: 151-175.