

AUTOCONCIENCIA E IDENTIDAD PERSONAL

MARTIN FRANCISCO FRICKE
UNAM, Instituto de Investigaciones Filosóficas

I. INTRODUCCIÓN

Supongamos que la identidad se exprese en oraciones de la forma $a = b$. Entonces podríamos decir que cada uno de nosotros tiene muchas identidades, las cuales puede expresar en oraciones donde el primer pronombre personal “yo” reemplaza la primera de las variables en el esquema mencionado. “Yo soy el primer hijo de tal y tal”. “Yo soy el cliente número un millón que compra en el supermercado tal y tal”. “Yo soy el investigador en filosofía que trabaja en lo que antes era la morgue del Sanatorio Rendón Peniche”. “Yo soy el ser humano que nació el día X en el lugar Y, con tal y tal historia.” Supongamos, además, que el pronombre “yo” en estas oraciones se refiera a una persona. Entonces podríamos decir que todas estas oraciones expresan las identidades de mi persona. ¿Podemos distinguir, entre estas identidades de mi persona, una que es *la* identidad personal que yo tengo? Quizá podamos decir que todas las identidades mencionadas forman diferentes aspectos de *la* identidad personal que yo tengo.

Ahora bien, en este trabajo investigo sólo ciertos aspectos de nuestra identidad personal, a saber, las identidades *temporales*. “Yo soy Martin F.” no expresa una identidad temporal porque sólo se refiere a una persona que está presente ahora (yo estoy presente ahora y también Martin F. está presente ahora). En cambio, “Yo soy el niño que se ve en esta foto” expresa una identidad diacrónica, una identidad entre una persona presente ahora (“yo”) y una persona que estaba presente en el pasado (“el niño”). La identidad diacrónica de una persona es resultado del hecho de que una persona persiste en el tiempo. Si una persona *persiste* por veinte años, es decir si vive por veinte años, entonces existen innumerables identidades temporales entre la persona en los diferentes momentos en que existió durante esa veintena de años de su vida. ¿Cuáles son las condiciones de esta persistencia de una persona en el tiempo? ¿Puede una persona permanecer en existencia aunque no esté consciente? ¿Puede una persona permanecer en existencia si ya no tiene memoria de su pasado? ¿Puede una persona permanecer en existencia aunque pierda su cuerpo?

Teorías neo-lockeanas de la identidad personal (diacrónica) generalmente enfatizan la importancia de la conciencia para el persistir de las personas a través del tiempo.¹ Según los neo-lockeanos, una persona no puede sobrevivir cambios muy drásticos en su psicología. La continuidad de la conciencia es necesaria para la persistencia de una persona. En el presente trabajo examino qué papel puede jugar la *autoconciencia* —un aspecto particular de la conciencia— en las condiciones de la persistencia personal. En especial me interesa preguntarme si es coherente afirmar —como lo han hecho algunos teóricos neo-lockeanos— que, en general, “la identidad de una persona no puede ser distinguida de lo que ella misma supone que es” (Rovane, 1990: 360). Argumentaré que ciertas concepciones de la relación entre autoconciencia e identidad personal son incoherentes o no plausibles y que las teorías animalistas de la identidad temporal de las personas están mejor fundadas que las teorías neo-lockeanas.

2. UN ARGUMENTO PARA LAS TEORÍAS NEO-LOCKEANAS

Antes de empezar con mis argumentos principales, déjenme especular un poco sobre las razones que nos invitan a considerar la autoconciencia como crucial para la identidad diacrónica de una persona. ¿Por qué se supone que la autoconciencia es importante en la persistencia de una persona en el tiempo?

Quizás el razonamiento sea el siguiente: cuando percibimos objetos materiales —entendidos como objetos con extensión, solidez y ubicación— siempre percibimos algunas de las propiedades que los hacen objetos *materiales*. Nuestra percepción nos revela dónde se encuentran (ubicación), los presenta como extendidos, con cierta superficie o con un peso específico. En cambio, al parecer podemos estar conscientes de nosotros mismos *sin* percibir nuestra extensión, solidez o ubicación. En apariencia es posible simplemente olvidarse de las propias propiedades materiales (por lo menos por un tiempo limitado), mientras se siga siendo consciente de sí mismo. “Estoy preocupado” podría ser una expresión de tal conciencia en la que estamos conscientes de un aspecto mental de nosotros mismos, pero quizá no de nuestra corporalidad. “He pensado en este problema

¹ Algunos teóricos neo-lockeanos son Derek Parfit, Sydney Shoemaker, Harold Noonan y Carol Rovane (véase la bibliografía). Todos se inspiran en John Locke, filósofo que sugirió que la identidad temporal de las personas depende de la memoria que cada persona tiene de su pasado. La formulación mejor conocida de su posición es la siguiente: “[‘Persona’] es, me parece, un ser pensante inteligente dotado de razón y reflexión, y que puede considerarse a sí mismo como el mismo, como una misma cosa pensante en diferentes tiempos y lugares; lo que tan sólo hace en virtud de su tener conciencia, que es algo inseparable del pensamiento y que, me parece, le es esencial, ya que es imposible que alguien perciba sin percibir que percibe” (Locke, Ensayo, libro II, cap. xxvii, §9 [p. 310].) Advierto que a continuación no trato de dar y discutir una correcta interpretación de la teoría de Locke. Más bien me interesa la coherencia y plausibilidad de tres concepciones específicas de la relación entre autoconciencia e identidad personal (véase inicio de sección 3). Estas tres concepciones se inspiran en Locke, pero no están pensadas como fieles interpretaciones de su teoría.

por varios días” podría ser otra conciencia de sí mismo que, al parecer, tampoco involucra necesariamente conciencia de la propia corporalidad o materialidad. Sin embargo, este último tipo de autoconciencia sí involucra una conciencia de la propia persistencia en el tiempo. Pensar en algo por varios días implica persistir a lo largo de ellos. Estar consciente de haber pensado en algo por varios días es una manera de estar consciente de la propia permanencia en el tiempo.

Ahora bien, esta diferencia entre autoconciencia y la percepción de objetos materiales podría invitar a las siguientes conclusiones: no podemos percibir objetos materiales sin percibir algunas de sus propiedades materiales. Por tanto, probablemente tales objetos no pueden persistir sin sus propiedades materiales. Pero dado que podemos estar conscientes de nuestra persistencia en el tiempo sin estar conscientes de nuestra materialidad como personas corporales, quizás esta materialidad sólo sea de segunda importancia para nuestra identidad temporal. ¿No podría pasar que nuestras propiedades materiales cambien, sean reemplazadas por otras o desaparezcan completamente mientras no estamos conscientes de *ellas* pero todavía seguimos siendo conscientes de nosotros *mismos*? Parece que el hecho de que tengamos o no ciertas propiedades materiales no importa para la manera en que estamos conscientes de nosotros mismos cuando nos olvidamos de nuestra propia corporalidad. Uno podría pensar que la persistencia propia como la conocemos en tal autoconciencia es lo único que cuenta para nuestra identidad temporal. Cualquier realidad material, mientras “produce”, en algún sentido, nuestra permanencia como la conocemos por nuestra autoconciencia, parece compatible con nuestra permanencia en el tiempo. En otras palabras, la identidad temporal como la conocemos por la autoconciencia podría parecer *esencial* para nuestra existencia temporal, mientras nuestras propiedades materiales podrían parecer *accidentales*.

Me parece que algún razonamiento de este tipo es la base de las teorías neolockeanas, según las cuales fenómenos “mentales” como la “continuidad” o “conectividad psicológica” son esenciales para la identidad temporal de las personas. Quiero hacer dos observaciones generales sobre este razonamiento:

(1) El razonamiento parece argüir por un dualismo de objetos materiales y personas: Personas son entidades radicalmente diferentes de objetos materiales. El dualismo probablemente no es cartesiano, ya que no identifica las personas con sustancias especiales (sino quizá más con universales). Pero el argumento para la diferencia entre objetos materiales y personas se parece al argumento cartesiano para el dualismo. Porque no percibo objetos materiales de la misma forma en que sé de la persona que soy, los dos han de tener diferentes condiciones de persistencia. Pero aquí vale notar que hay muchas concepciones del mundo que conciben objetos materiales en analogía con personas. Por ejemplo, Leibniz caracteriza los objetos materiales en analogía con la mente humana: ambos son mónadas. Creencias animistas conciben de objetos naturales como almas incorporadas. Kant no concibe objetos materiales en analogía con mentes, pero dice muy

claramente que los objetos materiales como son en sí podrían ser del mismo tipo que mentes consideradas *en sí* (cfr. *Crítica de la razón pura*, Paralogismos, A358). Nada de lo que sabemos excluye esta posibilidad. La existencia de estas concepciones de objetos materiales y de personas indica que la percepción que tenemos de ellos no necesariamente nos obliga a tener una concepción dualista. Sólo porque no percibimos no podemos concluir que las mentes de objetos materiales tales mentes no existen.

(2) El hecho de que quizás a veces no estemos conscientes de nuestras propiedades materiales no significa que podamos persistir sin ellas. Esta objeción es análoga a la objeción que se hizo al argumento cartesiano para la “distinción real” entre cuerpo y mente (cfr. Descartes, *Meditación Sexta*). Al parecer, Descartes dice: (a) No sé con certeza si tengo un cuerpo o no. (b) Sé con certeza que estoy pensando. (c) Por lo tanto, es posible que esté pensando sin tener un cuerpo. La objeción ya discutida en las objeciones y respuestas de la primera edición de las *Meditaciones* responde: No saber si *p* es verdadero, no es equivalente a saber que es posible que no *p*. El hecho de que Descartes no sepa si tiene un cuerpo no legitima la conclusión de que él podría existir sin cuerpo. Lo mismo ocurre en nuestro caso. El hecho de que, a veces, estamos conscientes de nuestra persistencia en el tiempo, por ejemplo al estar conscientes de haber pensado en una cosa por mucho tiempo, sin estar conscientes de nuestra materialidad, todavía no justifica la conclusión de que esta persistencia no sea dependiente de una materialidad muy específica.

Para ser más concluyente, el argumento cartesiano debe suponer que nuestra autoconciencia nos provee con un conocimiento de nuestra esencia. No sólo debo saber con certeza que estoy pensando, también debo saber que mi existencia sólo depende de este pensar; es decir, debo tener conocimiento de la esencia de la sustancia que soy como algo que sólo tiene el atributo de pensar y ningún atributo más. En el caso del neo-lockeano, el cual no argumenta a favor de la existencia de sustancias pensantes, mi autoconciencia debe proveerme con el conocimiento de “todo lo que es necesario” para que yo permanezca en existencia. Es decir, al estar consciente de haber pensado en algo por mucho tiempo sin estar consciente de la propia corporalidad, debo saber que estoy consciente de todo lo que es necesario para que yo siga existiendo, que no haya más que se deba saber para asegurarse de que yo permanecí en existencia. El neo-lockeano parece decir: Si estoy consciente, por ejemplo, de sufrir el mismo dolor como hace un rato, entonces ¿qué importa cuáles hechos materiales subyazcan a mi autoconciencia? ¿No es cierto que ya sé todo lo que se requiere para afirmar una identidad entre la persona que sintió el dolor hace un rato y la persona que ahora lo está sintiendo?

Antes de entrar a una discusión más detallada sólo quiero mencionar aquí que no es fácil para tal teoría aceptar que a veces nos equivocamos sobre nuestra identidad. Parece que estoy consciente de haber pensado en algo por muchos días o de sufrir el mismo dolor como hace un rato. Pero de hecho no he pensado

en el asunto por tantos días ni he sufrido el mismo dolor antes; simplemente tengo una conciencia errónea. Si tal error es posible, entonces ¿cómo puede ser la conciencia misma suficiente para nuestra identidad temporal y cómo puede ser independiente de los hechos materiales que la sostienen? ¿No necesitamos hechos fuera del ámbito de la conciencia “interna” para decidir si una persona persiste en el tiempo o no?

3. ¿ES POSIBLE QUE LA IDENTIDAD PERSONAL DEPENDA DE NUESTRA AUTOCONCIENCIA?

La afirmación de que la identidad diacrónica personal depende de nuestra autoconciencia se puede interpretar de varias formas. En lo que sigue, examino tres formas específicas: (1) Nuestra identidad temporal depende de un *conocimiento* que tenemos de ella, (2) Nuestra identidad temporal depende de la *creencia* de que tenemos tal identidad, (3) Nuestra identidad temporal depende de *decisiones* que hacemos con respecto a la identidad temporal que queremos poseer.

3.1 Autoconciencia como autoconocimiento

¿Podría el conocimiento de la propia identidad temporal ser condición de esta misma identidad? Me parece que tal condición es incoherente porque un conocimiento no puede ser condición de su propia verdad. La teoría sugiere que nuestra identidad temporal sólo se da si sabemos de ella. Eso significa que nuestro conocimiento de la propia identidad es una condición del hecho de que tenemos esta misma identidad, es decir un conocimiento es condición de su propia verdad. ¿Por qué no es posible que un conocimiento sea condición de su propia verdad? Supongamos que el conocimiento de que p sea condición necesaria del hecho de que p . Mi objeción contra esta concepción es que la adquisición del conocimiento de que p parece imposible. Antes de haber adquirido el conocimiento de que p todavía no es el caso de que p , porque sólo puede ser el caso de que p si ya existe el conocimiento de que p . De hecho, durante *todo* el proceso de adquisición del conocimiento todavía no puede ser verdad que p , porque mientras dure el proceso todavía no existe el conocimiento y por tanto aún no se da una condición necesaria del hecho de que p . Sólo cuando el proceso de la adquisición del conocimiento ha concluido y el conocimiento de que p ya existe puede ser verdad que p , porque ya se cumple la condición de que exista el conocimiento de este hecho. Pero eso significa, a mi parecer, que el conocimiento de que p nunca puede ser adquirido. ¿Por qué? Porque la adquisición de un conocimiento debe ser guiada, en algún sentido, por los hechos de los cuales tratará el conocimiento. Podemos ponerlo así: los posibles objetos del conocimiento deciden qué conocimiento podemos tener. La existencia del conocimiento no decide (no es lo que determina) qué objetos existen. Pero si el conocimiento de que p es condición

del hecho de que p entonces nunca se da este hecho durante la adquisición del conocimiento de él. Por tanto la adquisición del conocimiento de que p no puede ser guiada por el hecho de que p . Eso significa que no se puede adquirir tal conocimiento.²

Si este argumento es correcto, entonces el conocimiento de la propia identidad temporal no puede ser una condición de esta misma identidad —excepto si tuviéramos el conocimiento sin haberlo adquirido—. Pero la propia identidad en el tiempo es algo contingente y empírico. No parece plausible que podamos tener un conocimiento no adquirido de esta identidad. Ya que todo nuestro conocimiento de la propia identidad temporal debe haber sido adquirido, considero que este conocimiento no puede ser una condición de esta misma identidad. Nuestra identidad temporal no puede depender de nuestra autoconciencia si ésta se entiende como un conocimiento de la propia identidad temporal.

3.2 Autoconciencia como creencia acerca de sí mismo

¿Podría nuestra *creencia* de que tenemos cierta identidad temporal ser condición de esta misma identidad? Esta idea podría parecer como una solución de los problemas discutidos en la sección anterior. La formación de conocimiento debe conformarse con varias normas, pero quizá la formación de meras creencias sea menos restringida. Un conocimiento formado de manera incorrecta no es un conocimiento, pero una creencia formada de manera incorrecta todavía es una creencia. Sin embargo, me parece que la afirmación de que nuestra identidad temporal depende de que creamos que tenemos esta identidad es susceptible a objeciones similares, aunque menos fatales, como la primera posición examinada.

Para evaluar este tipo de teoría es interesante preguntarse qué pasa, según la teoría, con la identidad de una persona que cree que la teoría es verdadera. Supongamos que la identidad temporal de una persona dependa de que ella crea que tiene esta identidad temporal. ¿Qué pasa ahora si una persona descubre este hecho, si descubre que la forma de su identidad depende de que ella crea que la tiene? Me parece que en este caso la persona adquirirá nuevas capacidades de ase-

² Uno de los dictaminadores de este texto sugirió que aquí se debe distinguir entre “condición posibilitante” y “condición constitutiva”. A su entender, la condición posibilitante es un estado *previo* (“romper el cascarón es una condición necesaria previa para hacer un omelette”), mientras que la condición constitutiva es un estado *concomitante* (“el movimiento es una condición necesaria concomitante al batido de huevos”). El dictaminador sugiere que mi argumento contra la idea de que un conocimiento sea condición de su propia verdad depende de una noción posibilitante de condición. Me parece que su conclusión es errónea, pues incluso si la condición es constitutiva (y así concomitante) se da el problema anteriormente explicado: la adquisición del conocimiento en cuestión no puede ser guiada por el hecho del cual trata este conocimiento porque el hecho sólo se da si también se dan todas sus condiciones, sean éstas posibilitantes o concomitantes. No omito señalar que en Fricke 2002 he discutido tres posibles contraejemplos a mi tesis: la omnisciencia, los contratos y el dolor.

gurar u obstruir su propia identidad en el tiempo; capacidades que normalmente no nos atribuimos. Según la teoría, en cuanto podemos influir nuestras creencias sobre la propia identidad también podemos influir esta identidad misma. En el caso más simple una persona podría inducirse una amnesia que le hiciera perder todas sus creencias sobre su pasado. En consecuencia, la persona debería dejar de existir. La persona anterior habría cometido suicidio y, al mismo tiempo, habría creado una nueva persona (la persona amnésica) como resultado de inducirse la amnesia.

Incluso parece imaginable que una persona pudiera cambiar sus creencias sobre la propia identidad simplemente porque ella *quiere* cambiarlas. Supóngase que según la teoría, una persona A es idéntica con la persona X del pasado si cree que es idéntica con ella, o idéntica con la persona Y del pasado si cree que es idéntica con ésta. Ahora, si la persona A sabe que su creencia *hará* que ella sea idéntica con una de estas dos personas del pasado, parece que ella podrá *elegir qué creer* según sus preferencias. Si prefiere haber sido la persona X, simplemente cree que era la persona X antes; si prefiere haber sido la persona Y, simplemente cree que era la persona Y antes. Si prefiere no ser idéntica con ninguna de las dos —quizás ambas estén citadas en un juzgado por algún crimen— entonces simplemente cree que no es idéntica con ninguna de las dos. Según la teoría, sus creencias serán verdaderas en virtud de haber sido formadas.³

¿Pero es posible elegir de esta manera qué creer? Considerase en este contexto, la “lotería misteriosa” de Galen Strawson (1984: 207-211): Cualquier persona que cree que ganará en la lotería en realidad ganará. Aunque quizá haya un problema práctico en investigar quién sinceramente cree qué, parece que tal lotería es posible. El caso es análogo con la teoría propuesta sobre la identidad personal. La creencia de que uno ganará en la lotería es una condición necesaria (y en este caso también suficiente) de que uno ganará. Ahora supóngase que alguien descubra cuál es la condición para ganar en la lotería. Parece que esta persona puede llegar a formar la creencia de que ganará o que no ganará dependiendo de si *quiere* o no ganar:

[...] discovering that [...] believing you're a winner is necessary and sufficient for winning, and wanting to win, he can, surely, come to believe he will be a winner, knowing that his coming to believe this will make the belief true (he knows it's not just a matter of raising

³ Todo esto presupone que las creencias en cuestión no sean tácitas. Si fueran tácitas, es decir, creencias de las cuales no somos conscientes y las cuales no podemos articular, entonces no tendría mucho sentido decir que las *elegimos*. El que no está consciente de sus opciones no puede elegir entre ellas. La formación de creencias tácitas es automática en el sentido de que no ocurre a voluntad. Algunos piensan que la formación de creencias *siempre* es automática y nunca voluntaria, incluso cuando se trata de creencias conscientes. Si eso es correcto, entonces nunca es posible verdaderamente elegir las propias creencias. Como se mostrará más adelante, simpatizo con esta posición. Una persona racional cree porque las razones le obligan a creer, no porque elige hacerlo. Pero no siempre somos perfectamente racionales.

a hand, and that hand raisers will be tested for sincerity). But the wanting seems to be somehow necessary for the formation of the belief (Strawson, *op. cit.*: 210).⁴

Si aplicamos esta afirmación a la teoría de la identidad temporal de las personas que estamos considerando, entonces parece que alguien que quiere ser idéntico con una persona del pasado y no con otra y que sabe que su creencia sobre esta identidad es lo que determina esta cuestión, puede formar la creencia que quiere tener —y en este sentido puede elegir la forma de pasado que le conviene.

Sin embargo, Strawson nota —correctamente, en mi opinión— que hay “algo incómodamente falso y precario” en la formación de una creencia según las propias preferencias. Nos parece que la persona “en realidad no puede, o por lo menos no puede honestamente, formar [la creencia]” (*ibid.*). ¿Por qué tenemos esta impresión de falsedad y de engaño? Considero que la razón está en el hecho de que no es racional decidir creer algo de lo que uno *sabe* que no es verdadero —y esta es la situación en la cual se encuentra la persona si reflexiona suficientemente su formación de la creencia.

Déjenme explicar esta contención. Trivialmente, antes de haber formado una creencia acerca de que *p*, uno todavía no tiene una creencia acerca de que *p*. Es decir, antes de que el proceso de la formación de creencia haya sido completado, uno no tiene la creencia. Ahora bien, si es una condición del hecho de que *p* que uno crea que *p*, entonces no es verdad que *p* antes de que el proceso de formar la creencia de que *p* haya sido completado. De donde se sigue: si alguien sabe que es una condición del hecho de que *p* que él cree que *p* y si sabe que él todavía se encuentra en el proceso de formar la creencia de que *p*, entonces él puede concluir que (todavía) no es verdad que *p*. Este conocimiento hace que la formación de la creencia sea irracional. Saber que algo es falso y continuar aún así formando la creencia de que es verdadero es irracional. En cualquier momento de la formación de la creencia puedo saber que no es el caso de que *p*, porque en cualquier momento de dicha formación todavía no tengo la creencia de que *p* y por eso una condición del hecho de que *p* no está cumplida. Pero formar una creencia sabiendo que es falsa es irracional. En este sentido hay algo “incómodamente falso” y deshonesto en la formación de tal creencia. En algún momento la persona debe creer algo como lo siguiente: “Sé que no es verdad que *p*, pero lo voy a creer de todas maneras”. En el momento de creer que *p*, será verdad que *p*. Pero claramente, la decisión de creer que *p*, expresada en la oración “lo voy a creer de todas maneras” no constituye una manera racional de formar una creencia. Me parece que cuanto más reflexiva y racional sea una persona en

⁴ “[...] descubriendo que [...] creer que eres un ganador es necesario y suficiente para ganar y queriendo ganar, él puede, seguramente, llegar a creer que él será un ganador, sabiendo que su llegar a creer hará que la creencia será verdadera (él sabe que no es simplemente un asunto de levantar la mano y que los que levantan la mano serán examinados por su sinceridad). Pero el querer parece de alguna manera necesario para la formación de la creencia.”

formar sus creencias, menos probable es que adopte una creencia de la cual sabe que adoptarla es una condición de su verdad.

Si el argumento hasta aquí esgrimido es correcto, podemos distinguir dos consecuencias posibles en el caso de una persona que descubre que sus creencias acerca de la propia identidad temporal son necesarias para esta identidad:

(1) Si la persona no es muy racional y reflexiva en la formación de sus creencias, entonces, en circunstancias apropiadas, el descubrimiento le hace capaz de elegir con cuál persona del pasado (o del futuro) ella es idéntica o no idéntica, porque hasta cierto grado, ella será capaz de elegir qué creer sobre su pasado (o su futuro). Según la teoría, sus creencias, en torno, contribuyen y hacen posible que ella en realidad tenga la identidad temporal que se auto-adscribe.

(2) Sin embargo, si la persona es suficientemente racional y reflexiva entonces, al parecer, no será capaz de formar creencia alguna sobre su propia identidad temporal. Más bien se dará cuenta de que su decisión de creer que ella es idéntica con alguna persona del pasado o del futuro involucra decidir creer algo que (todavía) no es verdad y por tanto es una decisión irracional. Paradójicamente, mientras más racional sea la persona más corta será su vida porque no será capaz de formar las creencias acerca de la propia identidad que, según la teoría, son necesarias para que ella tenga esta identidad.

Me parece que nuestra conclusión debería ser que no es muy plausible afirmar que la identidad temporal de una persona depende de que ella crea que tenga esta identidad. El problema es que la teoría sugiere que las personas que son más reflexivas y racionales son por lo mismo también más vulnerables para rupturas de su identidad temporal. La teoría implica que cierta falta de conocimiento, reflexividad y racionalidad es *constitutiva* para la identidad temporal de las personas. Eso no significa que la teoría sea incoherente. A menudo encontramos una incompatibilidad similar entre la “meta-teoría” sobre algún fenómeno humano y la perspectiva de quienes están directamente involucrados en este fenómeno. Ejemplos de ello podrían ser discusiones de libertad, escepticismo, meta-ética o conocimiento. Sin embargo, aunque la incompatibilidad sea común, no es algo que incrementa la plausibilidad de una teoría. Además, simplemente parece extraño asumir que su racionalidad y conocimiento hacen que una persona sea menos capaz de persistir en el tiempo.⁵

⁵ Puede haber situaciones en las que es más probable que el individuo sobreviva si *no* usa su capacidad de razonar y reflexionar y en lugar de eso actúa de manera instintiva, por ejemplo huyendo de predadores (agradezco esta observación a uno de los dictaminadores). Lo que es extraño en la teoría que consideré en esta sección es que implica que la racionalidad, reflexividad y el conocimiento hacen que la persona sea menos capaz de persistir en el tiempo *incluso cuando la situación debería permitir el razonamiento y la reflexión*. Aun cuando no estamos amenazados por predadores sino que tenemos la oportunidad de reflexionar con tranquilidad la propia situación, la teoría parece implicar que tal reflexión racional hace que la persona sea menos capaz de persistir.

3.3 Algunas restricciones de teorías neo-lockeanas

En las dos secciones anteriores examiné y critiqué la idea de que nuestra identidad temporal podría depender de nuestro conocimiento de esta identidad o de nuestras creencias acerca de ella. La conclusión inmediata es que una teoría de la identidad temporal no debe postular que es una condición de nuestra identidad que sepamos de ella o que tengamos creencias acerca de ella. Pero quizá se pueden sacar unas conclusiones más fuertes de los argumentos desarrollados. Teorías neo-lockeanas de la identidad temporal de las personas dicen que nuestra identidad depende de la “conectividad” psicológica de la persona en el tiempo. Tal conectividad consiste en el hecho de que la persona recuerda su propio pasado, que tiene intenciones para acciones en el futuro y que sus conocimientos, creencias, otros estados intencionales, predilecciones, disposiciones, hábitos y rasgos de carácter, tienen cierta persistencia y continuidad en el tiempo. Los dos argumentos desarrollados anteriormente sugieren que no es coherente o plausible que la conectividad psicológica como es descrita por los neo-lockeanos incluya conocimiento o creencias acerca de la propia identidad temporal. Pero además, estos argumentos implican que todos aquellos tipos de conectividad psicológica que presuponen o involucran conocimiento o creencias acerca de la propia identidad temporal no pueden ser constitutivos de esta misma identidad. Permítaseme mencionar cuatro casos más específicos:

(1) La conectividad psicológica no puede incluir el conocimiento o la creencia de una persona de que ella ha *existido* en un tiempo anterior o que *existirá* en un tiempo posterior. Tales proposiciones no son equivalentes con proposiciones de identidad, pero las implican inmediatamente. Una persona no puede saber que ha existido en un tiempo anterior si es sólo gracias a este conocimiento suyo que ella es idéntica con la persona anterior (y por eso ha existido en este tiempo anterior). Tal conocimiento nunca podría ser adquirido. De igual manera, una persona no puede saber que existirá en un tiempo posterior si es sólo gracias a este conocimiento suyo que ella es idéntica con la persona posterior (y por eso existirá en este tiempo posterior).

(2) La conectividad psicológica no puede incluir muchos tipos de *intenciones*. Si uno tiene la intención de hacer algo en el futuro, la tiene con base en la creencia de que uno estará ahí —en el futuro— en alguna forma distinta. Alguien que cree que sólo tiene un año más de vida no tendrá planes para su futuro en diez años. Si formo la intención de hacer un millón de dólares en los próximos cinco años, debo creer que seguiré existiendo por los siguientes cinco años. Por eso no me parece plausible suponer que esta intención pudiera ser una condición de mi persistencia durante estos cinco años. Intenciones a corto plazo no son muy diferentes. Intentar hacer un café presupone algún tipo de creencia de que uno estará ahí para hacer el café. A menudo, intentar hacer algo requiere algún tipo de confianza de que el propio futuro permitirá realizar la intención y por eso no debería formar parte de la conec-

tividad psicológica, la cual es necesaria, según teorías lockeanas, para la identidad temporal de las personas.

El único tipo de intención que podría ser usado en las teorías lockeanas son las intenciones hipotéticas, tal como “Si sigo con vida mañana empezaré a ir a la iglesia”. Estas intenciones explícitamente cuentan con una incertidumbre sobre la propia existencia en el futuro, y, en tanto no presuponen creencias sobre esta existencia, no son vulnerables a los argumentos de las secciones anteriores.

(3) Esperanzas o temores y muchos otros estados emocionales como agitación, nerviosismo, tristeza o letargo pueden estar basados en conocimiento o creencia acerca del propio futuro o acerca de lo que uno ha experimentado antes. Uno sólo puede estar nervioso por una cita con el dentista, por ejemplo, si uno cree que uno existirá en el futuro (y realizará la cita). Un prisionero teme el día cuando, según lo que cree, tendrá que compartir la celda con otro prisionero violento. Sin embargo, si se entera de que, en una fecha previa, será ejecutado, ya no tendrá el mismo miedo, sino otro. Alguien podría estar deprimido sobre algo malo que cree que ha hecho en el pasado. Descubriendo que la creencia es falsa, que la maldad pasó antes de que él naciera, la depresión probablemente se aliviará.

Estos ejemplos muestran que muchos estados emocionales dependen de conocimiento o creencias acerca de la propia existencia en el pasado o en el futuro. De ello se sigue que, para evitar los problemas señalados en las secciones anteriores, estos estados emocionales no deberían ser incluidos en la noción de conectividad psicológica como la usan las teorías neo-lockeanas de la identidad personal.

(4) Si los *recuerdos* del propio pasado se consideran como una forma de conocimiento del propio pasado, entonces deben también ser excluidos de la noción de conectividad psicológica usada en las teorías neo-lockeanas. Porque al ser incluidos, el conocimiento del propio pasado —en la forma de recuerdos— sería al mismo tiempo una condición de este pasado como propio. Es decir, un conocimiento sería condición de su propia verdad. Uno no tendría un pasado si uno no supiera que lo tiene. Mi argumento anterior pregunta cómo, en este caso, uno puede adquirir este conocimiento del propio pasado, dado que este pasado no es propio hasta que uno ya sepa que lo es. Si recuerdos del propio pasado —entendidos como conocimiento del pasado— fueran una condición de ese pasado, entonces nuestra memoria no tendría nada de recordar.

Sin embargo, quizás el caso de la memoria es un poco más complicado. Podemos acaso tener recuerdos del pasado sin que éstos constituyeran conocimiento del propio pasado. Podría ser que la función de la memoria sea una facultad simulativa que nos permite re-vivir un evento del pasado y así adquirir conocimiento del pasado. En este caso, la memoria no sería conocimiento sino una capacidad para simular acontecimientos del pasado, en el sentido de crear experiencias similares a las del pasado. La memoria misma no sería conocimiento del pasado sino sólo una posible base para tal conocimiento. Para que el neo-lockeano pueda usar

la noción de memoria como constitutiva de la identidad personal, debe postular que la siguiente idea tenga sentido: una persona se pregunta cuáles eventos del pasado puede recordar. Pero no recordar en el sentido ordinario en el que se supone que tal recordar es acordarse del propio pasado; más bien, la persona tiene que preguntarse de manera más “neutral” cuáles acontecimientos del pasado puede simular o re-vivir. Con base en su posibilidad de re-vivir (predominantemente) acontecimientos que ocurrieron en la vida de determinada persona, se puede concluir, según la teoría neo-lockeana, que esta persona es la misma que la que ahora está simulando esas experiencias. Me parece evidente que tal razonamiento está muy lejos de nuestras experiencias ordinarias. Normalmente nuestros recuerdos se ven de inmediato acompañados por la creencia de que se trata de nuestro pasado.

Si el neo-lockeano acepta estas cuatro restricciones de la noción de conectividad psicológica, entonces sólo le queda un número limitado de fenómenos psicológicos para explicar la identidad temporal de las personas. Éstos podrían incluir los siguientes: estados intencionales que no presuponen conocimiento o creencias acerca de la propia identidad y existencia en otros tiempos, preferencias, disposiciones, rasgos de carácter, estados emocionales que no dependen de conocimiento o creencias acerca de la propia identidad, intenciones hipotéticas y simulaciones del pasado (o del futuro, si eso es posible) que pueden existir independientemente de conocimiento o creencias acerca de la propia existencia en el pasado (o en el futuro). Aquí hay unos ejemplos de los fenómenos mentales que podrían conformar la conectividad psicológica según el neo-lockeano: la creencia de que los teléfonos celulares dañan la salud, la aracnofobia (fobia a las arañas), la preferencia por cierta comida, rasgos de carácter como paciencia, irritabilidad o valentía... La conectividad psicológica, y con ella la identidad temporal, entonces, dependería de la preservación y del desarrollo coherente de estos fenómenos en el tiempo. Según Quassim Cassam, una motivación importante de las teorías neo-lockeanas es explicar “la identidad personal en términos que corresponden al sentido que la persona tiene de ser la misma cosa pensante en diferentes tiempos” (Cassam, 1997: 178). Si mi crítica es correcta, entonces hay restricciones importantes para el intento de realizar esta motivación. La identidad temporal de una persona, por tanto, no puede depender de su propio conocimiento o su propia creencia acerca de esta misma identidad y tampoco de estados mentales que a su vez dependen de tal conocimiento o de tales creencias.

3.4 Autoconciencia como agencia

Quizás exista otro tipo de teoría de la identidad temporal de las personas que puede satisfacer las motivaciones de los neo-lockeanos. Personas no sólo son seres que *saben* y *creen* algo de sí mismos. Personas también son *agentes* que pueden ejecutar acciones según su propia voluntad. Estas acciones tienen un impacto

sobre su identidad y persistencia en el tiempo. Comemos, trabajamos, dormimos, nos cuidamos —sin estas acciones no podríamos persistir en el tiempo. Una persona también puede suicidarse o vivir de una manera riesgosa y así tener un impacto negativo en su persistencia en el tiempo.

Ahora bien, supóngase que uno pudiera tomar posesión de su identidad pasada o de su identidad en el futuro de una manera similar a un colono que toma posesión de un lote de tierra todavía no ocupado. Quizá, dadas ciertas circunstancias, uno puede tener cierta identidad temporal simplemente por *declarar* la identidad suya. Yo soy la persona que hizo tal y tal cosa en el pasado porque yo declaro que soy esta persona, pero si no tomo posesión de esta identidad, la identidad no es mía y no soy la persona que hizo tal y tal cosa en el pasado. Si esto fuera una posibilidad real, entonces la identidad temporal de una persona dependería de sus propias acciones en un sentido mucho más fuerte que el considerado anteriormente. Las acciones más importantes para la identidad temporal de una persona serían, entonces, aquellas en las que ella toma posesión de su propia identidad. La persona sería capaz de formar su identidad en el tiempo decidiendo de cuál identidad toma la posesión.⁶

Nótese que este tipo de teoría no padece de los problemas anteriormente desarrollados. No hay incoherencia en la idea de que las acciones de una persona influyen en la identidad temporal de esta persona. Claro está que una persona ya tiene que existir, en algún sentido, para ejecutar cualquier acción, pero eso no significa que la acción no podría tener un impacto en la identidad temporal de la persona. Probablemente se sigue que una persona puede tener una identidad temporal aun sin realizar la acción de tomar posesión de alguna identidad. Sin embargo, esto es compatible con la idea de que la identidad de las personas *también* depende de si o no una persona realice ciertas acciones en cierto momento. Incluso si suponemos que la persona misma sabe del impacto que sus acciones pueden tener en su propia identidad y si suponemos que la persona es reflexiva y racional, no surgen los problemas de plausibilidad discutidos antes. El colono sabe que tomar posesión de un lote de tierra hará que el lote se torne suyo. Su conocimiento de esta condición de poseer un lote no dificulta la acción de tomar posesión. Si fuera posible realizar acciones similares con respecto a la propia identidad temporal, entonces no tendríamos el problema que la acción involucrara irracionalidad en personas que saben que sus acciones determinan

⁶ En Fricke 1999 propuse una teoría semejante. En lo que sigue explicaré por qué ahora creo que no es coherente con nuestro concepto de identidad. Robert Nozick elabora una teoría que comparte algunos rasgos con la idea que aquí discuto. Según él, la identidad temporal de una persona podría depender de los criterios que la persona misma acepta. Así, personas que aceptan diferentes teorías de identidad personal, en consecuencia, también tienen diferentes identidades temporales. *Ceteris paribus*, mi supervivencia depende de la teoría de la supervivencia que tengo (cf. Nozick, 1981, cap. 1, II.).

su propia identidad. En ningún momento sería necesario que la persona formase una creencia sabiendo que ésta (todavía) no es verdadera.

En el marco de una teoría de este tipo, se interpretarían las diferencias entre las concepciones de la identidad temporal de las personas en diferentes culturas, religiones o convicciones personales de una manera novedosa: no es necesario suponer que la mayoría de las concepciones sean falsas. Más bien se podría tratar de diferentes opciones que las personas pueden realizar en las elecciones de su identidad personal. Las discusiones en la filosofía contemporánea podrían ser interpretadas similarmente: las intuiciones contrarias que provocan los distintos casos imaginarios (trasplantes de cerebros, transferencias de memorias, división o fusión de personas, etc.) podrían expresar diferentes preferencias en lugar de opiniones verdaderas o falsas. Nuestra actitud frente a todas estas concepciones de la identidad temporal de las personas no debería enfocarse en el intento de descubrir cuál concepción es la correcta, sino en averiguar cuál es la que mejor corresponde a nuestras preferencias y, por tanto, debería ser elegida por y para nosotros.

Los problemas de una teoría del tipo esbozado son lógicos. La teoría implica (1) que la identidad no es transitiva y (2) que la verdad de enunciados sobre la identidad es relativa al tiempo en que se hace el enunciado. Además, (3) la teoría dificulta concebir personas como entidades que persisten en el tiempo.

Nada en la teoría, como la he esbozado, excluye la posibilidad de que dos personas distintas “tomen posesión” de identidades temporales que se “cruzan” en ciertos momentos. Supongamos que la persona A decide que ella es idéntica con una persona X en el futuro (o en el pasado) y supongamos que la persona B también decide que ella es idéntica con la persona X. Entonces la persona A es idéntica con la persona X y la persona B también es idéntica con la persona X, pero no hay duda de que las personas A y B *no* son idénticas, porque se trata de dos personas distintas en el presente. Tal constelación contradice claramente la idea de que la identidad es una relación transitiva: si el autor de la *Crítica de la razón pura* es idéntico con Immanuel Kant, y el autor de los *Prolegómenos* es idéntico con Immanuel Kant, entonces el autor de la *Crítica de la razón pura* es idéntico con el autor de los *Prolegómenos*.

Que la teoría hace la identidad relativa al tiempo se ve cuando consideramos la posibilidad de que una persona cambie de opinión acerca de su identidad temporal. Supongamos que la persona A decide que ella es idéntica con cierta persona B en el futuro y supongamos que la situación es tal que su mera decisión de ser así idéntica hace que la persona A sea idéntica con la persona B. Parece que nada en la teoría excluye la posibilidad de que la persona A cambie su decisión en un momento posterior. (Podemos suponer que no haya duda —ni siquiera en su propia mente— de que se trate de la misma persona A en ambos momentos.) Quizá la persona A descubre algo desagradable en la persona B y por eso ahora decide que no es idéntica con la futura persona B sino con otra futura persona, la persona C. En consecuencia, según la teoría ya no debería ser el caso que la

persona A sea idéntica con la persona B; más bien debería ser idéntica con la persona C. Eso significa que la identidad es relativa al tiempo. Si la persona A es idéntica con la persona B o no, depende del momento en que se hace la pregunta. ¿Pero cómo es eso posible? Parece que o dos cosas son idénticas (numéricamente), una y la misma cosa, o no lo son. Pero si son idénticas entonces siempre es verdad que son idénticas y si no son idénticas entonces siempre es verdad que no lo son. El autor de la *Crítica de la razón pura* es idéntico con el autor de los *Prolegómenos*, no importa cuándo investigue el asunto. Parece que el tiempo no puede cambiar si dos cosas en realidad son una sola, la identidad no parece relativa al tiempo.

Consideremos el fenómeno de la persistencia en el tiempo. Si una cosa, por ejemplo una persona, ha persistido en el tiempo hasta el presente, entonces eso es lo que ha pasado a la cosa. Parece que no es posible aislar la cosa como está en el presente y cambiar el hecho de que ha persistido en el tiempo en cierta manera y que, gracias a esta persistencia, es idéntica con cierta cosa que ha existido en el pasado. Lo que le ha pasado a la cosa necesariamente pertenece a la cosa presente. Afirmar que uno podría elegir un pasado diferente para la cosa es negar que ella haya persistido desde un tiempo anterior hasta el presente. La persistencia del pasado ya ha “pasado” y por eso ya no se puede cambiar en el presente. Podemos decidir y cambiar lo que pasará en el futuro. Sin embargo, estas decisiones y estos cambios conciernen a cosas (incluso personas) que persistirán en el futuro. Podemos, hasta cierto grado, decidir si una cosa persistirá en el futuro y en qué forma lo hará y estas decisiones se pueden cambiar más tarde. Pero tal cambio de decisión sólo significa que la decisión anterior no fue realizada; no significa que antes de la reversión de la decisión la cosa tuviese una identidad temporal diferente. Una vez establecido que una cosa persiste en cierta forma en el futuro, no existe una cuestión *adicional* acerca de la identidad temporal de la cosa en su futuro. La identidad temporal implica persistencia y viceversa. No existe la posibilidad de una “distribución arbitraria de identidades” sobre las cosas persistentes; no es posible que el simple “tomar posesión” de una identidad temporal decida cuál es la identidad de una persona.

4. DOS MANERAS DE ACERCARSE AL PROBLEMA DE LA IDENTIDAD TEMPORAL DE LAS PERSONAS

Hemos investigado tres versiones de la idea de que la autoconciencia es condición de la identidad temporal de las personas. La primera versión —que sugiere que un conocimiento de la propia identidad temporal es condición de esta identidad— resultó incoherente: El conocimiento en cuestión no puede ser adquirido. La segunda versión —según la cual una creencia sobre la forma de la propia identidad temporal es condición de esta identidad— resultó no plausible: Una persona que cree que la teoría es verdadera y que reflexiona la propia formación de creencias sobre la identidad temporal ya no puede formar las creencias

necesarias para esta misma identidad. La tercera versión —que hace la identidad dependiente de un “tomar posesión” de la identidad— resultó incoherente con la lógica de nuestra noción de identidad. En conclusión, no es fácil construir una teoría neo-lockeana de la identidad temporal de las personas si se postula que, en general, “la identidad de una persona no puede ser distinguida de lo que ella misma supone que es” (Rovane, *op. cit.*: 360).

Para terminar, quiero distinguir entre dos diferentes clases de razones que pueden motivar una teoría de la identidad temporal de las personas. Me parece que esta distinción nos puede ayudar a evaluar diferentes estrategias en la discusión de la identidad personal. ¿Por qué investigamos la identidad temporal de las personas? ¿Por qué nos interesa esta materia? Me parece que existen dos tipos de razones. La primera clase de razones proviene de ciertos desacuerdos e incertidumbres que se derivan muy naturalmente de nuestra vida ordinaria. Por ejemplo, quizá no tengamos claridad acerca de cuándo una persona empieza a existir y, por tanto, hasta qué momento deberíamos permitir el aborto y si es permitido experimentar con embriones tempranos. También puede pasar que no tengamos certeza acerca de cuándo una persona deja de existir, por ejemplo cuando discutimos la muerte clínica y la posibilidad de donar órganos o la pregunta de si es ético apagar las máquinas que mantienen a una persona comatosa con vida. Además es posible que tengamos dudas acerca del impacto que ciertas condiciones psicológicas como la llamada personalidad múltiple tienen sobre nuestra identidad personal. Otros desacuerdos pueden surgir de la confrontación de diferentes religiones y de éstas con perspectivas no religiosas o ateas. Entre todos estos sistemas de creencia existen muchos desacuerdos acerca de qué son personas y bajo cuáles condiciones persisten.⁷ Finalmente, hay numerosos experimentos mentales que involucran teletransportación, trasplantes de cabeza o de cerebro, transferencias de memorias e intenciones, duplicaciones o fusiones de personas y otras ideas ficticias, las cuales tienden a provocar juicios divergentes y contradictorios sobre la identidad temporal de las personas cuando diferentes sujetos las examinan. Todos estos desacuerdos, los ordinarios, los interreligiosos y los derivados de la imaginación, parecen requerir alguna forma de solución y, por tanto, nos dan razones para investigar el problema de la identidad temporal de las personas con el objetivo de llegar a alguna teoría que pueda arbitrar entre estos desacuerdos.

⁷ En muchos de los casos anteriores se mezclan dos problemas: el de la identidad temporal y el de la constitución de las personas. (Debo esta observación a un árbitro.) Si alguien dice “Yo, alguna vez, fui este feto” (señalando a la impresión de un escaneo ultrasonográfico), entonces hace una afirmación sobre su identidad temporal (cf. Olson 2010, sección 2). Otra persona quizás negaría tal afirmación porque cree que ella misma, como persona, todavía no existía cuando ya existía el feto. Es posible que en este caso el desacuerdo no sólo sea acerca de las condiciones de persistencia de una persona (identidad diacrónica) sino también sobre qué nos permite llamar a algo o alguien una persona.

El segundo tipo de razones para investigar la identidad personal es más simple. Se encuentra en el hecho obvio de que existen personas en el mundo y que es innegable que ellas persisten en el tiempo. Es posible que haya muchos casos problemáticos en los cuales es debatible si se dan estos hechos, pero también hay casos (ordinarios) en los cuales no se puede dudar que sí se dan. Dados estos casos no controversiales, uno simplemente puede preguntarse qué subyace y explica tal persistencia de las personas por el tiempo. ¿Gracias a qué es el caso que en estas circunstancias ordinarias una persona persista en el tiempo?

Evidentemente, los dos tipos de razón para investigar la identidad temporal de las personas están relacionados. Por un lado, tratar de descubrir qué subyace en la persistencia de las personas en circunstancias no controversiales podría llevarnos a una teoría que nos ayude a arbitrar en los casos más problemáticos. Por otro lado, tratar de solucionar los acuerdos “naturales” sobre la identidad temporal que surgen de la vida ordinaria nos llevará a una teoría que también tiene algo que decir sobre los casos no controversiales de la identidad temporal. Sin embargo, me parece que los dos tipos de razones normalmente tienden a motivar diferentes tipos de teorías de la identidad personal. Si empezamos por exponer y discutir los casos problemáticos y controversiales de la identidad, es probable que la teoría final se base en consideraciones de plausibilidad intuitiva asociada con experimentos mentales más y más remotos. La teoría entonces describe facetas de nuestro concepto *intuitivo* y *no reflexionado* de la identidad temporal de las personas. Al concentrarnos en algunos aspectos de este concepto, podemos tratar de revisarlo, hacerlo más preciso y llegar a una decisión sobre los casos controversiales con los que empezamos. El problema con este acercamiento al fenómeno de la identidad temporal de las personas es, en mi opinión, que es difícil justificar la intuición particular (o el conjunto de tales intuiciones) en la que se basa nuestra revisión del concepto de la identidad.

Si, por otro lado, nuestra teoría está motivada por el simple hecho de que personas existen y persisten en el tiempo y por la suposición de que debe haber algo que subyazca a este hecho, entonces, me parece, es menos probable que nos enredemos en una batalla de intuiciones no reflexionadas. La teoría entonces se concentrará menos en lo que pensamos y sentimos que somos y, en lugar de eso, enfocará en lo que está ahí. Es posible, e incluso probable, que la teoría llegará a la conclusión de que lo que está “ahí”, cuando hablamos de personas y su persistencia, es un poco diferente de lo que ordinariamente pensamos que esté ahí, que las entidades cuya existencia subyace el hecho de que existen y persisten personas son diferentes de lo que ordinariamente concebimos como personas. Sin embargo, aquí este aspecto revisionista de la teoría no está motivado por simples intuiciones no reflexionadas. Más bien, está basado en la idea de que haya hechos que *subyacen* en nuestra percepción ordinaria del mundo y de nosotros mismos. Estos hechos están relacionados con y, en algún sentido, son responsa-

bles de nuestra percepción, pero no necesariamente son lo que pensamos es el caso antes de investigar lo que subyace nuestra percepción.⁸

El candidato principal para la entidad que subyace casos no controversiales de la identidad temporal de las personas es —me parece— el *animal humano* que, en estos casos, persiste en el tiempo. El animal humano es la entidad a la que naturalmente podemos atribuir todos los aspectos de la persistencia ordinaria de personas como nosotros: Tiene un cuerpo que se modifica en el tiempo y tiene una vida psicológica que se desarrolla según la historia que el animal vive en el mundo. Si tenemos que revisar y definir el concepto de persona según los hechos que subyacen la existencia y persistencia ordinaria de personas como nosotros, me parece que este concepto debe ser considerado como el concepto de un animal de algún tipo, un organismo animal humano, que se convierte en un mero cuerpo cuando muere. Las propiedades psicológicas sólo juegan un papel menor en este tipo de teoría de la identidad personal. Es difícil hallar una entidad substancial que corresponda, como su poseedor, exclusivamente a la psicología que tenemos, una entidad que exclusivamente posea nuestra psicología pero no todas las otras propiedades que tiene el animal humano que somos. En casos ordinarios de la persistencia personal en el tiempo, tanto las propiedades psicológicas como las corporales se adscriben de la manera más natural al animal humano. Cuando un animal humano persiste en el tiempo, sus órganos persisten, ciertas células persisten y ciertas moléculas y átomos persisten. Uno podría tratar de aislar una de esas entidades y adscribirle un estatus especial, declarando que *esto* es la persona. Sin embargo, me parece que, al menos en los casos no controversiales de la persistencia personal, tal definición del concepto de persona no estaría bien justificada. Un concepto de persona así definido separaría artificialmente algunas partes de la entidad sobresaliente: el animal humano. En los casos ordinarios, es indudablemente el animal humano completo el que persiste y sostiene la mayoría de los atributos que adscribimos a personas.⁹

La idea de que personas como nosotros son animales humanos es, entonces, el punto de partida para una teoría de la identidad temporal de las personas

⁸ Uno de los dictaminadores señaló, acertadamente, que en este segundo caso también nos basamos en intuiciones; la única diferencia parece ser que se trata de intuiciones diferentes, quizá de corte biológico. Pero creo que hay otra: en el primer caso, partiendo de situaciones problemáticas en la vida cotidiana, *sólo* nos basamos en intuiciones, que no tienen un respaldo adicional. La investigación se limita a medir el “peso” de cada intuición. En el segundo caso tenemos cierto marco conceptual; investigamos qué es lo que subyace a los fenómenos ordinarios. Aquí las intuiciones en las que nos basamos son principalmente de corte ontológico. Nos preguntamos qué “cosas” realmente están ahí. Quizá también podamos guiarnos en lo que se reconoce como entidades reales en las ciencias. Tenemos, entonces, cierto criterio para decidir cuáles intuiciones son relevantes. Sólo lo son aquellas que tratan de qué está ahí (y, quizá, que se integren coherentemente con lo que ya sabemos sobre este asunto).

⁹ Entre los autores que recientemente han defendido teorías animalistas de la identidad personal se incluyen Paul Snowdon, Eric Olson y David Mackie (véase la bibliografía).

que es motivada por el deseo de investigar qué subyace en la persistencia ordinaria de las personas. Con este punto de partida, la teoría puede tratar de solucionar los casos más complejos de la identidad temporal, que por lo general producen desacuerdos. Probablemente las soluciones estarán en conflicto con algunas de nuestras intuiciones existentes sobre estos casos. En particular, la teoría animalista contradirá teorías neo-lockeanas de la identidad personal según las cuales las propiedades psicológicas son el factor más importante en la persistencia de las personas. Como traté de mostrar, tales teorías normalmente están basadas en las intuiciones sobre experimentos mentales y casos problemáticos de la identidad temporal. Me parece que las teorías animalistas pueden contradecir las teorías neo-lockeanas con buenas razones: los hechos que subyacen en los casos normales de persistencia personal simplemente no apoyan la idea de que otra entidad que no sea el animal completo corresponda a lo que llamamos persona. En este sentido el concepto animalista de persona está bien fundado. Eso no es decir que sea la mejor “reconstrucción” posible del concepto de persona. Pero la teoría animalista presenta un reto fuerte contra las teorías basadas en simples intuiciones sobre la identidad temporal de las personas: ¿Por qué deberíamos aceptar estas intuiciones? ¿Por qué deberíamos usarlas como el elemento determinante de nuestro concepto de persona? ¿Por qué deberíamos revisar nuestro concepto a la luz de lo que dicen estas intuiciones y no en vista de lo que parece evidente de casos no problemáticos de persistencia personal? Quizá las intuiciones en cuestión puedan ser justificadas —entre los límites que prescriben la coherencia, la plausibilidad y la lógica de nuestro concepto de identidad que he desarrollado anteriormente— y quizá se puedan encontrar buenas razones para suponer que personas no necesariamente son animales, sino que puedan persistir aunque el animal no lo haga. Sin embargo, para justificar estas intuiciones se necesitarían argumentos. No basta con sólo señalar que existen.¹⁰

¹⁰ Una versión anterior de este texto fue presentada como ponencia en el VI Coloquio de Epistemología “Identidad y memoria” celebrado en Mérida en 2006. Agradezco los comentarios de la audiencia en esa ocasión, así como los comentarios de los dos dictaminadores anónimos de *Península*. Estoy particularmente en deuda con Paul Snowdon por las observaciones que hizo a la primera versión de este trabajo durante mis estudios de doctorado. La labor en la versión aquí publicada se benefició del apoyo del proyecto “La primera persona y sus percepciones” (PAPIIT IN400508).

BIBLIOGRAFÍA

- CASSAM, Quassim
1997 *Self and World*. Oxford, Oxford University Press.
- DESCARTES, René
1975 *Meditaciones metafísicas*. Buenos Aires, Aguilar.
- FRICKE, Martin
1999 “Selbstbewusstsein und personale Identität”. *Rationalität, Realismus, Revision*, pp. 493-500, Julian Nida-Rümelin (ed.). Berlin & New York, Walter de Gruyter.
2002 *A Phenomenological Theory of Self-Consciousness*. Thesis (DPhil), Oxford, Oxford University.
- KANT, Immanuel
2006 *Crítica de la razón pura*. México, Taurus.
- LOCKE, John
1999 *Ensayo sobre el entendimiento humano*. México, FCE.
- MACKIE, David
1999 “Personal Identity and Dead People”, *Philosophical Studies* 95 (3): 219-242.
- NOONAN, Harold
2003 *Personal Identity*. London, Routledge. 2a ed.
- NOZICK, Robert
1981 *Philosophical Explanations*. Oxford, Clarendon Press.
- OLSON, Eric T.
1997 *The Human Animal: Personal Identity without Psychology*. Oxford, Oxford University Press.
2010 “Personal Identity”, *The Stanford Encyclopedia of Philosophy (Winter 2010 Edition)*, Edward N. Zalta (ed.). Consultado en <http://plato.stanford.edu/archives/win2010/entries/identity-personal/> el 30 de enero de 2010.
- PARFIT, Derek
2004 *Razones y personas*. Madrid, Machado
- ROVANE, Carol
1990 “Branching Self-Consciousness”, *Philosophical Review* 99 (3): 355-396.
- SHOEMAKER, Sydney
1984 “Personal Identity: A Materialist Account”, *Personal Identity*, pp. 67-132, Sydney Shoemaker y Richard Swinburne (eds.). Oxford, Blackwell.
- SNOWDON, Paul
1990 “Persons, Animals and Ourselves”, *The Person and the Human Mind*, pp. 83-107, Christopher Gill (ed.). Oxford, Clarendon Press.
- STRAWSON, Galen
1984 *Freedom and Belief*. Oxford, Oxford University Press.