



Revista Colombiana de Filosofía de la  
Ciencia

ISSN: 0124-4620

revistafilosofiaciencia@unbosque.edu.co

Universidad El Bosque  
Colombia

Rivera Novoa, Angel  
Dilemas morales, juicio moral y corteza prefrontal ventromedial  
Revista Colombiana de Filosofía de la Ciencia, vol. 13, núm. 27, julio-diciembre, 2013, pp.  
43-61  
Universidad El Bosque  
Bogotá, Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=41431644002>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica  
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

# DILEMAS MORALES, JUICIO MORAL Y CORTEZA PREFRONTAL VENTROMEDIAL<sup>1,2</sup>

## MORAL DILEMMAS, MORAL JUDGMENT AND VENTROMEDIAL PREFRONTAL CORTEX

Angel Rivera Novoa<sup>3</sup>

### RESUMEN

La investigación empírica acerca de fenómenos neurológicos se ha convertido, en los últimos años, en un elemento relevante a la hora de analizar conceptos morales. Este artículo analiza la función de la corteza prefrontal ventromedial en la producción de juicios morales. Se analizan dos investigaciones a pacientes con daño en la corteza y se argumenta que la interpretación de los resultados no es adecuada. Se propone, por un lado, una diferencia entre dilemas personales e impersonales con base en la posibilidad de utilizar a una persona como medio y, por otro, que la anormalidad en la respuesta a los dilemas, en pacientes con daño en la corteza, no es una opción por el utilitarismo, sino una respuesta azarosa. Con estos dos elementos, la interpretación de los resultados se hace consistente.

**Palabras clave:** dilemas morales personales, dilemas morales impersonales, utilitarismo, emociones, daño en la corteza prefrontal ventromedial.

### ABSTRACT

Empirical research on neurological phenomena has become, in recent years, a relevant point when we analyze moral concepts. This article analyzes the role of the ventromedial prefrontal cortex in the production of moral judgments. Two investigations with patients with damage to the cortex are analyzed and it is argued that the interpretation of the results is not appropriate. This paper proposes, first, a difference between personal and impersonal dilemmas based on the possibility of using a person as a means. Secondly, it is proposed that the abnormality in response to dilemmas in patients with damage to the cortex is not an option for utilitarianism, but a random answer. With these two elements, the interpretation of the results is consistent.

**Keywords:** moral personal dilemmas, moral impersonal dilemmas, utilitarianism, emotions, damage to ventromedial prefrontal cortex.

1 Recibido: 21 de septiembre de 2013. Aceptado: 13 de noviembre de 2013.

2 Agradezco al profesor Alejandro Rosas y a los participantes del seminario de Ética Analítica del segundo semestre de 2010 en la Universidad Nacional de Colombia por los comentarios hechos a una versión previa de este escrito. En el artículo Rosas, Caviades, Arciniegas, Arciniegas (2013) se presentan conclusiones similares a las de este artículo, pero, por un lado, el modo de llegar a ellas es diferente y, por otro, no se utilizan elementos del mencionado artículo en éste, por lo que no se tomará como punto de referencia.

3 Universidad Nacional de Colombia. Correo electrónico: angelrivera32@gmail.com

En los últimos años, ha aumentado notablemente la investigación empírica acerca de los fenómenos neurológicos, psicológicos y biológicos subyacentes a la moral. En particular, el rol causal que podrían tener las emociones en la producción de los juicios morales ha sido centro de análisis. Green *et al.* (2001, 2004), por ejemplo, han desarrollado ciertos experimentos en los cuales se muestra que hay una activación de la corteza prefrontal ventromedial (en adelante CPV) - región del cerebro asociada a las emociones - cuando los sujetos se enfrentan a un tipo de dilemas morales particulares.

Green *et al.* distinguen dos clases de dilemas morales: personales e impersonales. Los dilemas morales personales son aquellos en los cuales se propone aceptar (o no) realizar un daño físico directo a una persona a cambio de un logro o un resultado de la acción de bienestar general. El ejemplo paradigmático de este tipo de dilemas es el dilema del puente, que consiste en empujar a un hombre gordo de un puente para que detenga un tren que, de no ser detenido, mataría a cinco personas ubicadas en los rieles. Por otra parte, los dilemas morales impersonales son aquellos en los que se causaría un daño colateral a una persona (esto es, sin intención), a cambio de un resultado de la acción también de bienestar general. El ejemplo más conocido y ampliamente discutido es el dilema del tren, cuyo escenario consiste en un tren que, igualmente, se dirige a matar a cinco personas. Una persona tiene la posibilidad de halar una palanca u oprimir un botón para desviar el rumbo del tren; sin embargo, esto traería como consecuencia que el tren arrollaría a una persona en vez de cinco.

Una buena manera para desarrollar la investigación es estudiar casos de pacientes con daño en la CPV (en adelante pacientes CPV) y analizar su comportamiento ante dilemas morales personales. Si la hipótesis según la cual la CPV tiene un rol importante en la producción normal de juicios morales es correcta, entonces los pacientes CPV deberían mostrar una anormalidad a la hora de emitir juicios ante dilemas morales personales. En este ensayo, analizaré dos experimentos que van este línea (*cf.* Ciaramelli *et al.* y Koenigs *et al.*). Mi propósito es mostrar ciertas inconsistencias entre los resultados obtenidos y el análisis que hacen los investigadores sobre los mismos. Luego de analizar tales problemas, me propongo hacer una lectura diferente de los resultados obtenidos, por medio de una distinción más clara entre los dos tipos de dilemas arriba expuestos, con el fin de eliminar los problemas mencionados para apoyar la hipótesis de que la CPV cumple un rol causal en la producción de juicios morales de algún tipo particular.

Para tal propósito, primero, expondré la diferencia entre dilemas morales personales e impersonales, así como la diferencia entre dilemas morales personales

difíciles y fáciles. En segundo lugar, analizaré el experimento de Ciaramelli *et al.*, señalando tres problemas. A continuación, analizaré el experimento de Koenigs *et al.*, mostrando dos problemas adicionales. Por último, propongo reinterpretar los criterios de la distinción de los dilemas, así como lo que debe entenderse por ‘respuesta anormal’ de pacientes CPV a dilemas morales personales, con lo cual pretendo disolver los problemas de los experimentos. Con esto se favorece la hipótesis de que la CPV juega un papel causal en el juicio moral.

## 1. LOS EXPERIMENTOS DE GREEN *ET AL.* Y LOS DILEMAS MORALES

Green *et al.* (2001) quieren mostrar que, desde un punto de vista psicológico, la activación emocional es un elemento fundamental del juicio moral. En particular, quieren mostrar que en los casos de dilemas morales personales, la activación de las emociones es crucial para el juicio moral —de ahí que, en el ejemplo del puente, la mayoría de las personas no acepta la acción propuesta (sólo 11% de aceptación)—. No ocurre lo mismo, en cambio, en el caso de los dilemas morales impersonales —razón por la cual, la mayoría de las personas aprueban la acción propuesta (89% de aceptación)— (*cf.* Green *et al.* 2001, 2105-2106). Green *et al.* enfocan su experimento hacia el hecho de que las áreas del cerebro asociadas a las emociones deben mostrar una activación cuando un individuo contempla un escenario moral personal, pero no en el caso de los dilemas morales impersonales.

Cuando la acción propuesta de un dilema moral personal es aprobada, el tiempo de respuesta debe ser notablemente mayor a cuando se reprueba, lo que mostraría la influencia de mecanismos cognitivo-rationales que chocarían con la activación emocional a la hora de elegir (*cf.* Green *et al.* 2004, 390).

¿Cuál es entonces la explicación de la diferencia entre la respuesta al dilema del puente y la respuesta al dilema del tren? Por un lado, hay una diferencia entre la configuración misma de los dilemas (morales personales y morales impersonales) y, por otro, hay una activación cerebral asociada a las emociones en el caso de los dilemas personales, pero no en el caso de los impersonales. La diferencia según Green *et al.* (2001, 2004) entre unos y otros dilemas consiste en que, en los dilemas personales, las violaciones morales supuestas deben cumplir en conjunto tres condiciones: 1) causar un daño corporal severo, 2) el daño debe estar dirigido a una persona o conjunto de personas y 3) el daño no debe ser colateral, esto es, se debe suponer una agencia o una intención. Por el contrario los dilemas que no suponen la reunión de estas tres condiciones son impersonales (*cf.* Green *et al.* 2004, 389).

A su vez, hay una diferencia entre dilemas morales personales difíciles y dilemas morales personales fáciles. Según Green *et al.* (2004), las respuestas a los dilemas morales personales difíciles son lentas y no presentan una uniformidad. Esto se debe a que la respuesta emocional negativa, asociada a la acción propuesta en el dilema, entra en conflicto con un razonamiento abstracto de maximización del bienestar, necesario según la situación. Un ejemplo de este tipo de dilemas es el del llanto del bebé, según el cual, unos militares enemigos llegan a un pueblo. La persona, con muchas otras, se esconde en una casa y tiene en sus brazos un bebé. Cuando los militares están cerca, el bebé comienza a llorar fuertemente. La persona tapa la boca del niño. El dilema consiste en, o bien quitar la mano de la boca del niño, lo que causaría que los militares los descubrieran y los mataran a todos, o bien no quitar la mano, de tal modo que no serían descubiertos, pero tendría que asfixiar al bebé.

Los dilemas morales personales fáciles, en contraste, suponen respuestas rápidas y uniformes, ya que la respuesta emocional negativa, asociada al pensamiento de la acción propuesta, domina fácilmente a la débil o incluso inexistente deliberación racional a favor de la acción. Un caso típico de estos dilemas es el del infanticidio, según el cual una madre debe decidir si mata o no a su bebé recién nacido, el cual ha sido no deseado.

Quiero resaltar, debido a que esto será importante más adelante, que en ambos casos las emociones juegan un papel crucial a la hora de emitir un juicio moral. En el caso de los dilemas morales personales difíciles, las emociones juegan un papel en su lucha con el razonamiento. En el caso de los dilemas morales personales fáciles, las emociones son fundamentales para las respuestas dadas que, como muestran Green *et al.* (2004), son rápidas y uniformes.

En los experimentos de Green *et al.* (2001, 2004), entonces, se le presentaban este tipo de dilemas a los sujetos y se les indagaba sobre si aprobarían la acción repulsiva o no. Al momento de responder los dilemas, los sujetos fueron introducidos en una máquina de resonancia magnética que permite capturar, en imágenes, ciertas activaciones cerebrales. Los experimentos mostraron que, al menos en el caso de los dilemas morales personales, hay activaciones en las zonas del cerebro asociadas a las emociones. No obstante, no es claro, dados estos experimentos, si tales activaciones son causa o consecuencia de los juicios morales. Así, los experimentos no son concluyentes para el objetivo en cuestión, a saber, establecer si las emociones tienen una conexión causal con los juicios morales. Pasaré entonces a revisar los experimentos de Ciaramelli *et al.* y de Koenigs *et al.*, los cuales intentan directamente mostrar que las emociones sí cumplen un rol causal y no sólo colateral en los juicios morales (al menos personales).

## 2. INTENCIÓN SECUNDARIA Y EL EXPERIMENTO

### DE CIARAMELLI *ET AL*

En el experimento de Ciaramelli *et al.*, se ponen a consideración 45 dilemas de tres tipos diferentes: dilemas no-morales, dilemas morales impersonales y dilemas morales personales (15 de cada clase y todos de la misma batería de los de Green *et al.* 2001). Se seleccionaron dos grupos de personas: 7 pacientes CPV y 12 sujetos normales (estos últimos para hacer control del experimento). Los pacientes CPV mostraban déficit en su conducta personal social, carencia de seguimiento de normas sociales, así como una conducta emocional disminuida. Tanto los pacientes como los sujetos se sentaban frente a un computador que les iba mostrando dilemas de los tres tipos mencionados. La presentación de los dilemas tenía tres partes. En un primer momento, se presentaba el escenario del dilema; luego, se preguntaba a los sujetos si aprobarían o desaprobaban la acción propuesta; por último, se hacía una pregunta acerca del escenario para verificar si había una alteración de memoria a la hora de deliberar qué respuesta se daba. La predicción de los investigadores fue la siguiente:

Si las regiones mediales prefrontales están implicadas en el rechazo a violaciones morales personales, entonces los pacientes con lesiones en esta región deben estar más inclinados que los saludables-control [sujetos normales] a aprobar violaciones morales en dilemas morales personales. En contraste, ninguna diferencia en el resultado fue esperada entre pacientes y [sujetos] control en dilemas morales impersonales y dilemas no-morales, en los cuales la conducta es considerada como menos dependiente de los procesos en las áreas mediales prefrontales. (Ciaramelli *et al.* 85)

Algunos resultados fueron los siguientes: en general, el tiempo en que tanto pacientes como sujetos normales aprobaron la acción en dilemas morales (tanto personales como impersonales) fue mayor que el de la aprobación de las acciones en dilemas no-morales. A su vez, en general se dieron menos respuestas aprobatorias en los casos de dilemas morales (tanto personales como impersonales) en contraposición a los dilemas no-morales.

Los resultados más interesantes parecen haberse dado en el contraste entre dilemas morales personales e impersonales. El experimento mostró que los sujetos normales fueron más lentos en dar la respuesta a dilemas morales personales que al dar respuesta a los impersonales. En cambio, los pacientes CPV tuvieron un promedio de tiempo de respuesta muy similar en ambos casos. Además, los sujetos normales eran rápidos para desaprobatar las acciones propuestas en los dilemas morales personales y lentos en los casos de aproba-

ción —aunque el tiempo de respuesta de aprobado o desaprobado fue similar en casos de dilemas morales impersonales—. Por su parte, los pacientes CPV mostraron tiempos similares de respuesta entre aprobado o desaprobado tanto en dilemas personales como en impersonales. Más aún, los pacientes CPV fueron mucho más rápidos que los pacientes normales en aprobar acciones en dilemas morales personales, aunque no en el caso de dilemas impersonales.

En general, el número de respuestas aprobatorias fue menor en el caso de los dilemas morales personales. Los sujetos normales dieron menos respuestas aprobatorias a los dilemas personales que a los impersonales. Los pacientes CPV, en cambio, mantuvieron una proporción similar al aprobar acciones en escenarios tanto personales como impersonales. Por último, el número de respuestas aprobatorias en escenarios personales por parte de los pacientes CPV fue mayor que el de los sujetos normales. En cambio, las respuestas aprobatorias en dilemas morales impersonales tuvieron el mismo rango en ambos grupos.

Las pacientes CPV, en resumen, fueron más rápidos y más inclinados que los sujetos normales a autorizar violaciones en dilemas morales personales, pero sus respuestas y el tiempo de las mismas fue similar en el caso de dilemas morales impersonales. En contraste, los sujetos normales estuvieron menos inclinados a aprobar violaciones personales en comparación con las violaciones impersonales, mientras que los pacientes CPV mostraron una conducta similar en ambos escenarios.

Lo anterior permitiría concluir que la CPV juega un rol importante en la producción de los juicios morales personales. En efecto, dado que los sujetos normales fueron reacios al aceptar violaciones morales en escenarios de dilemas personales y, en cambio, los pacientes CPV estuvieron inclinados a una aprobación más numerosa y más rápida, entonces la CPV debe influir en tales juicios. Así, dado el vínculo entre la CPV y ciertas emociones, se concluye que las emociones son necesarias para la producción normal de juicios morales personales.

¿Cómo se da el vínculo entre la CPV y los juicios morales? Según Ciaramelli et al., en el curso de una decisión, los sujetos toman en cuenta tanto las consecuencias inmediatas como aquellas que son a largo plazo. La CPV permitiría la anticipación de emociones negativas tales como la culpa o la vergüenza cuando se piensa en causarle daño al individuo, lo cual evitaría llevar a cabo violaciones morales. Por tal razón, los pacientes CPV presentarían una inclinación a aprobar violaciones morales personales, debido a que, al tener lesiones en la CPV, no serían capaces de anticipar emociones negativas a largo plazo que los inhibieran de aceptar la violación propuesta.

Ahora bien, ¿por qué la anticipación de emociones negativas no se aplica también a casos impersonales? ¿Es legítimo suponer que en los casos impersonales también es posible anticipar ciertas emociones de culpa o emociones similares? De ser así, habría una incongruencia con los resultados dados en el experimento, pues, en tal caso, los sujetos normales, al anticipar emociones negativas, tenderían a rechazar de igual modo la violación propuesta en los casos de dilemas morales impersonales.

Pero, ¿por qué podríamos hacernos la pregunta sobre el caso de los dilemas morales impersonales? Al comparar el dilema del tren con el del puente, vemos que, esquemáticamente, el resultado es el mismo: se evita la muerte de cinco personas, pero el costo de ello es la vida de uno. ¿Acaso ser responsable de la vida de una persona no es suficiente para la anticipación de emociones negativas? Entonces, ¿por qué no hay un rechazo sistemático de este tipo de acciones propuestas en escenarios impersonales como se da en los personales? Alguien podría decir que esto es imposible, pues como muestran los experimentos de Green *et al.* (2001, 2004) —que son los que proveen la batería de dilemas—, ante los escenarios morales impersonales no hay una activación de la CPV. Si no hay una activación de la CPV, es imposible la anticipación de emociones negativas. No obstante, al preguntarse por qué no hay un rechazo sistemático de la acción propuesta en estos dilemas, lo que al mismo tiempo se pregunta es por qué justamente no hay una activación de la CPV si es posible pensar que tales escenarios generen emociones negativas.

Ciaramelli *et al.* anticipan este punto señalando:

Como Lieberman (2006) recientemente ha apuntado, los dilemas personales pueden inducir a los individuos a centrarse en su propia participación personal en dar lugar a un resultado desagradable, mientras los dilemas impersonales, por definición, carecen de este sentido de agencia y responsabilidad. (Ciaramelli *et al.* 90)

En últimas, lo que se dice es que, aunque el daño causado sea el mismo en ambos escenarios —personales e impersonales—, en un caso hay agencia y en el otro no. Esta diferencia daría la explicación de por qué en un caso se rechaza sistemáticamente la acción propuesta y en el otro no. No obstante, creo que este punto no es lo suficientemente claro. Evidentemente hay una diferencia entre causar un daño directo a una persona (en el caso de los dilemas personales) y causarle un daño indirecto (en los dilemas impersonales). Pero puede argumentarse fácilmente que la diferencia, en términos de responsabilidad, no es cualitativa sino cuantitativa, esto es, no necesariamente se debe suponer, como apuntan los investigadores, que en los dilemas los agentes carecen de agencia y responsabilidad totalmente.

En efecto, creo que si fuera así, entonces el agente que oprime el botón para desviar el tren no sería tampoco responsable de salvar la vida a las cinco personas, así como no sería responsable de quitar la vida a otro. Es posible hacer una diferencia entre responsabilidad o agencia primaria y secundaria. Una persona puede sentirse responsable en dos situaciones diferentes: uno de esos casos es cuando se causa un daño intencional a otro agente —*v.g.* una persona llevada por la ira golpea repetidas veces a otra persona hasta dejarlo inconsciente—. Pero este no es el único tipo de eventos en donde un agente puede sentirse responsable; puede ser el caso que, aunque no haya intención de lastimar a alguien, esto de todas formas ocurra —*v.g.* en un descuido, una persona deja una información a la vista que lastima innecesariamente a un tercero—; en este caso, la persona que deja la información puede sentirse responsable aunque no haya tenido la intención de lastimar. En este último tipo de casos, hay un sentido de agencia y responsabilidad secundaria. La diferencia entre la agencia/responsabilidad primaria y secundaria es que en el primer caso no hay una intención de hacer daño, en el segundo sí, pero en todo caso se genera un daño a un tercero. Este daño, sea causado intencionalmente o no, puede generar un sentimiento de culpa y, por tanto, un sentido de agencia y responsabilidad.

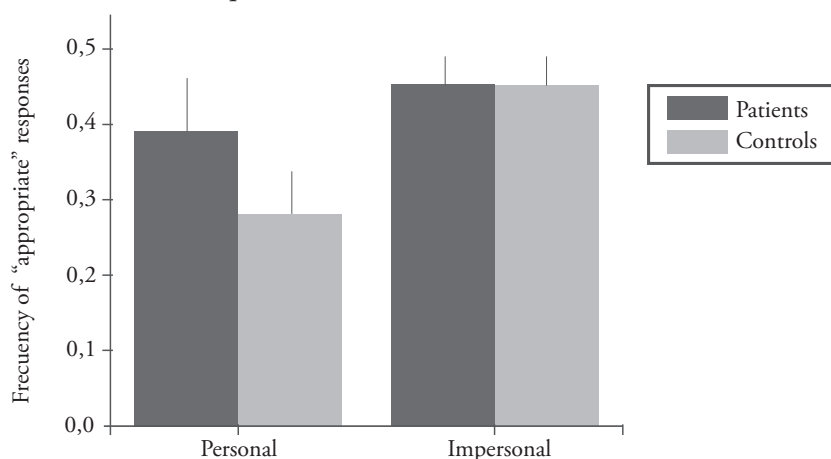
En el caso de los dilemas impersonales, habría una responsabilidad secundaria que debe tomarse en cuenta a la hora de tomar la decisión o de aprobar o desaprobar la acción propuesta. Pero si esto es así, no hay nada que impida afirmar que esa responsabilidad secundaria es relevante para anticipar emociones negativas. Por ejemplo, quien oprime el botón podría considerar que en un futuro sentiría culpa, pues su acción de oprimir el botón podría causar la muerte de una persona aunque, como supone el dilema, no sería debido a un daño directo. Si esto es así, no se explica por qué la tasa de aprobación en el caso de dilemas impersonales en sujetos normales no es más baja. Si la emoción anticipada en este caso diera lugar, los resultados deberían ser similares en el caso de dilemas personales e impersonales.

Alguien podría argumentar que si la emoción anticipada es proporcional a la intención de daño, entonces ésta se activaría sólo en el caso en el que hay una intención primaria de daño, pero no cuando la intención es secundaria. Pero la emoción anticipada puede también ser proporcional al resultado de la acción. En últimas, no hay evidencia, de entrada, para afirmar que la emoción anticipada es proporcional a la intención y no al daño generado. En el caso del dilema del tren, el resultado es la muerte de una persona, lo cual es obviamente indeseable. Tal resultado indeseable anticiparía una emoción que provocaría la desaprobación de la acción propuesta. Pero esto va en contra de los resultados obtenidos.

En segundo lugar, creo que si es correcto afirmar que hay una diferencia entre una agencia o responsabilidad directa y otra secundaria, entonces la distinción entre los dilemas morales personales y los dilemas morales impersonales se hace oscura. En efecto, tal como se mostró anteriormente, los dilemas morales personales deben cumplir tres condiciones, a saber, 1) causar un fuerte daño físico 2) a un ser humano 3) a través de una agencia. Pero entonces los dilemas impersonales podrían muy bien cumplir con las tres condiciones, pues involucrarían una agencia secundaria proporcional al resultado indeseable.

Bien podría argumentarse que la condición 3 debe suponer una agencia o responsabilidad directa. Sin embargo, no veo evidencia relevante para colocar esta condición, de nuevo, porque creo que la diferencia entre una agencia primaria y otra secundaria es apenas de grado. Más aún, tampoco veo evidencia suficiente para suponer que las emociones que surgirían en uno u otro escenario sean radicalmente diferentes. El dilema del tren y el del puente supone sacrificar la vida de alguien y, sea como sea el sacrificio, el resultado es claramente indeseable.

Quiero señalar un tercer problema también en relación a los resultados de los experimentos. Si bien hay una diferencia entre las respuestas de los pacientes CPV y los sujetos normales, en el caso de los dilemas morales personales (0.39 de aprobación en pacientes CPV contra 0.28 en sujetos normales sobre una escala de 1), tal diferencia no parece significativa en relación a la hipótesis del experimento (*cf.* Ciaramelli *et al.* 88, fig.3). En efecto, si se supone que los pacientes CPV presentan un déficit selectivo, esto es, una respuesta anormal de aprobación en los juicios morales personales, el porcentaje de aprobación resulta muy bajo. Si los pacientes CPV sólo centraran su atención en la utilidad de la aprobación de la acción, entonces la tasa de aprobación debería estar cercana a 1, cosa que no ocurre:



Se podría argumentar que la tasa de aprobación no es cercana a 1, debido a que los dilemas morales personales tipo Green contienen dos categorías, a saber, los difíciles y los fáciles (cf. Green 2004, 390-391). La tasa no sería tan alta, debido a que un número de dilemas presentados a los pacientes CPV serían dilemas personales fáciles. No obstante, creo que esto no sería un buen argumento, debido a que, como ya lo había advertido, en todo caso las emociones jugarían un papel importante en ambas categorías. En los dilemas difíciles chocarían con las consideraciones racionales y en los fáciles implicarían una rápida desaprobación de la acción. Los pacientes CPV, al presentar un déficit emocional selectivo, de igual forma tendrían que presentar una tasa alta de aprobación en todo dilema moral personal.

En resumen, no es claro cómo la CPV anticipa emociones negativas, porque esto iría en contra de los resultados de los experimentos, ya que los sujetos normales deberían mostrar una menor tasa de aprobación en los dilemas morales impersonales. Asimismo, debe establecerse un criterio más claro para diferenciar dilemas morales personales de los impersonales, pues la diferencia, en estos experimentos radica, en parte, en la agencia. Pero como mostré, es posible pensar una agencia secundaria. Por último, uno esperaría una tasa más alta de aprobación por parte de los pacientes CPV en el caso de los dilemas morales personales, si tienen un déficit selectivo. Al finalizar este ensayo, mostraré cómo estos resultados pueden ser coherentes.

### 3. CONFIGURACIÓN DE DILEMAS Y EL EXPERIMENTO DE KOENIGS *ET AL.*

Koenigs *et al.* desarrollan un experimento – muy similar al de Ciaramelli *et al.* – que involucra a seis pacientes CPV, sujetos normales y pacientes con daños en otras zonas del cerebro (en adelante, pacientes tipo 2) y dilemas morales personales, impersonales y no-morales. Tal experimento muestra que los pacientes CPV exhiben un patrón utilitarista anormal de juicios sobre dilemas morales determinados. Los dilemas puestos a consideración, de nuevo, contraponen consideraciones convincentes de bienestar en conjunto contra conductas altas y emocionalmente repulsivas. Los juicios de los pacientes CPV, en dilemas morales impersonales y no-morales, fueron normales. Así, para un conjunto específico de dilemas morales – los personales –, la CPV es esencial para los juicios normales, de tal modo que las emociones cumplirían un papel necesario en la generación de estos juicios.

La aprobación de la acción propuesta – supone el modelo de Koenigs *et al.* – requiere que el sujeto supere una respuesta emocional contra el daño *directo*

hacia otra persona, lo que genera un juicio utilitarista que maximiza el bienestar en cada situación. La hipótesis planteada dice que si las respuestas emocionales relacionadas con la CPV tienen un rol causal en los juicios morales, entonces los pacientes CPV mostrarían un índice alto anormal de juicios utilitaristas en escenarios morales personales. Sin embargo, los pacientes CPV deberían presentar un patrón normal de juicios morales impersonales y de juicios no-morales. Si las emociones no juegan un papel crucial en la producción de juicios morales, los pacientes CPV deberían mostrar un patrón normal en todos los escenarios presentados.

El experimento mostró que no hubo diferencias esenciales entre los tres grupos ante escenarios de dilemas no-morales y morales impersonales. Pero en los escenarios de dilemas morales personales, los pacientes CPV fueron mucho más propensos a aceptar la acción propuesta que los dos grupos restantes. A su vez, no hubo diferencia entre el grupo de sujetos normales y los pacientes tipo 2. Esto muestra que las respuestas del grupo de pacientes CPV tienen una alteración sólo en los juicios sobre dilemas morales personales, lo que sugiere que la mediación de la CPV afecta sólo aquellos juicios morales que involucren acciones emocionalmente significativas (*cf.* Koenigs *et al.* 1909).

Se analizó más a fondo el conjunto de respuestas dadas en los dilemas morales personales, lo cual llevó a los investigadores a realizar una diferencia entre escenarios de bajo conflicto y escenarios de alto conflicto — aparentemente análoga a la diferencia de Green *et al.* (2001, 2004) entre dilemas morales personales fáciles y difíciles —. Siete de los veintidós escenarios de dilemas personales presentaron un acuerdo en relación a las respuestas de los tres grupos. Otro dilema adicional, presentó un acuerdo casi total entre los tres grupos. Fue en virtud de los trece escenarios restantes que hubo una notable diferencia entre el grupo de pacientes CPV y los otros dos grupos. Los ocho primeros escenarios fueron calificados entonces de bajo conflicto — todos los pacientes CPV rechazaron la acción propuesta —, mientras que los trece restantes fueron clasificados como de alto conflicto. Todos los escenarios de alto conflicto provocaron que un gran número de pacientes CPV aceptaran la acción propuesta, en contraste con los otros dos grupos. El tiempo de reacción ante el escenario, dicen los investigadores, también habla a favor de la distinción (en el grupo de sujetos normales, por ejemplo, el tiempo de respuesta en los escenarios de alto conflicto fue notablemente más alto que en el resto de los escenarios).

En resumen, los pacientes CPV mostraron un patrón de respuesta utilitarista anormalmente alto en comparación con el grupo de pacientes normales y con pacientes tipo 2. Así, las respuestas de los pacientes CPV fueron producto

de un cálculo utilitarista en virtud de maximizar el bienestar general, ante escenarios morales personales de alto conflicto con ausencia de una reacción emocional al potencial daño directo hacia otra persona. De esta manera, el experimento concluye que las emociones juegan un papel importante en la producción de juicios morales personales de alto conflicto, debido al rol causal de la CPV en la producción de los mismos.

Quiero señalar, en primer lugar, una duda acerca de la estrategia metodológica que asumen los investigadores acerca de la distinción entre escenarios de conflicto alto y bajo. Tal como se planteó la hipótesis del experimento, se quería probar si la CPV cumple un rol causal en la producción normal de juicios morales personales (sin más). Además, las respuestas al resto de dilemas debían presentar un patrón normal. Sin embargo, tal y como Koenigs *et al.* presentan el experimento, no hubo, de hecho, un patrón anormal en la respuesta esperada, pues, como mostré, en una gran cantidad de escenarios morales personales hubo una respuesta normal, igual a la de los otros dos grupos de pacientes. A partir de esto, los investigadores hacen la diferencia entre escenarios de conflicto alto y bajo. Lo que parece extraño es que la diferencia no fue dada en la formulación de la hipótesis, sino sólo a raíz de los resultados. ¿Acaso esta metodología no podría ser calificada como *ad hoc*? Si fuese así, el resultado general desmentiría la hipótesis inicial ya que, de hecho, no hubo un patrón anormal utilitarista de respuesta en pacientes CPV, sino sólo en algunos de los escenarios propuestos. No obstante, se podría argüir entonces que, justamente, la diferencia entre los escenarios se traza a partir de los resultados. Green *et al.* (2004, 390-391) distinguen los dilemas personales difíciles de los fáciles a partir del patrón de respuesta.

Quiero entonces evaluar la estructura de algunos de los dilemas usados por Koenigs *et al.*, para mostrar que algunos de ellos presentan una configuración extraña que viciaría los resultados y, por ende, la tesis de los investigadores. Veamos el siguiente dilema que se califica como dilema moral personal de bajo conflicto:

You are driving along a country road when you hear a plea for help coming from some roadside bushes. You pull over and encounter a man whose legs are covered with blood. The man explains that he has had an accident while hiking and asks you to take him to a nearby hospital. Your initial inclination is to help this man, who will probably lose his leg if he does not get to the hospital soon. However, if you give this man a lift, his blood will ruin the leather upholstery of your car. Would you leave this man by the side of the road in order to preserve your leather upholstery? (Koenigs *et al.* 2007. Supplementary Information: 11)

Creo que este dilema no ejemplifica correctamente el modelo que deben seguir los dilemas personales (sean fáciles o difíciles). Los dilemas de este tipo suponían la elección entre una opción utilitarista de bienestar general contra conductas reprochables. Este dilema, y otros con la misma estructura, no exhiben la posibilidad de esta elección.

El no manchar el auto con la sangre del herido no representa en ningún caso una opción de bienestar general utilitarista (no es, por ejemplo, análogo a salvar cinco vidas). Por ende, no constituye ningún dilema real, con base en el cual pueda analizarse el déficit selectivo, o no, de los pacientes CPV. Contrataré ahora dos dilemas. El primero de ellos es el dilema del puente, el cual es calificado como un escenario de conflicto alto. El otro, es el famoso dilema del médico, que es calificado como un escenario de conflicto bajo:

A runaway trolley is heading down the tracks toward five workmen who will be killed if the trolley proceeds on its present course. You are on a footbridge over the tracks, in between the approaching trolley and the five workmen. Next to you on this footbridge is a stranger who happens to be very large. The only way to save the lives of the five workmen is to push this stranger off the bridge and into the tracks below where his large body will stop the trolley. The stranger will die if you do this, but the five workmen will be saved. Would you push the stranger on to the tracks in order to save the five workmen?

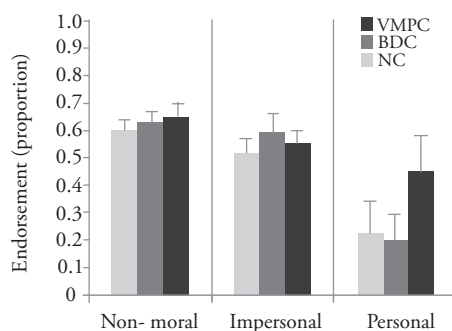
[...] You are a doctor. You have five patients, each of whom is about to die due to a failing organ of some kind. You have another patient who is healthy. The only way that you can save the lives of the first five patients is to transplant five of this young man's organs (against his will) into the bodies of the other five patients. If you do this, the young man will die, but the other five patients will live. Would you perform this transplant in order to save five of your patients? (Koenigs *et al.* 2007. Supplementary Information: 11)

Las situaciones, a pesar de ser muy diferentes, exhiben una estructura idéntica. Se hace un terrible daño a una persona, esto es, se produce su muerte de un modo intencional, en virtud de un resultado utilitarista, a saber, salvar la vida de cinco personas. Sin embargo, el primero de estos dilemas se clasifica como de alto conflicto y el otro de bajo conflicto. ¿Por qué si exhiben la misma estructura?

La diferencia se hace con base a la respuesta, pero, al tener una estructura idéntica, y ser clasificados de forma diferente, los resultados nuevamente estarían viciados. La diferencia entre ambos tipos de dilemas (de alto y bajo conflicto) puede entonces sí ser *ad hoc*, pues la diferencia es artificial y totalmente dependiente de las respuestas. Dada la configuración de los dilemas

(unos no exhiben dilemas reales y otros se clasifican de modo distinto con una estructura idéntica), los resultados no parecen conclusivos.<sup>4</sup>

Otro problema que presenta el experimento de Koenigs *et al.* es análogo a uno de los problemas que señalé en relación al experimento de Ciaramelli *et al.*: hay una baja respuesta de aprobación, por parte de los pacientes CPV, en relación a los juicios morales personales. En efecto, aunque en tales escenarios hay una diferencia entre las respuestas de los pacientes CPV y los sujetos normales junto con los pacientes tipo 2, de nuevo, tal diferencia no parece ser una diferencia significativa a la luz de la hipótesis. En una escala de 1, los pacientes CPV aprobaron la acción en escenarios morales personales en un promedio que no llega ni siquiera a 0.5. Como en el caso de Ciaramelli *et al.*, se esperaría que el rango de aprobación de la acción propuesta en los escenarios morales personales fuera mucho más alta. Se esperaría que estuviera cercana a 1 (cf. Koenigs *et al.* 910):



**Figure 2. Moral judgements for each scenario type.** Proportions of “yes” judgements are shown for each subject group: Error bars indicate 95% confidence intervals. We used three classes of stimuli: Non-moral scenarios ( $n = 18$ ), impersonal moral scenarios ( $n = 11$ ), and personal moral scenarios ( $n = 21$ ). On personal moral scenarios, the frequency of endorsing “yes” responses was significantly greater in the VMPC group than in either comparison group ( $p$  values  $< 0.05$ , corrected).

Koenigs *et al.* podrían argüir que la tasa de aceptación no es tan alta debido a la diferencia entre escenarios de bajo y alto conflicto. Sin embargo, como en el caso de Ciaramelli *et al.*, esto no sería un buen argumento. Las emociones,

<sup>4</sup> Aunque Kahane y Shackel (2008) no muestran uno a uno un análisis de algún dilema particular como los que acabo de presentar, tales autores señalan que los escenarios propuestos por Koenigs *et al.* no muestran escenarios de dilemas reales (tal como mostré que ocurría con el hombre que no quería ensuciar su auto). Según estos autores, sólo un 45% de los dilemas impersonales y un 48% de los dilemas personales son de hecho dilemas. También señalan que la diferencia entre escenarios de alto y bajo conflicto no corresponde a una diferencia en el contenido de los escenarios mismos (como mostré en el contraste de los dilemas del puente y del médico) y la diferencia sólo corresponde a una diferencia en la conducta de respuesta — lo que es consistente con mi argumento — (cf. Kahane y Chackel).

como ya lo expuse anteriormente, según la división de Green *et al.*, juegan un papel importante en ambos tipos de escenario. En uno entran en conflicto con el razonamiento y, en otro, dominan completamente el razonamiento. De esta manera, el resultado esperado sería una aprobación utilitarista anormal en todos los escenarios morales personales. Por otra parte, las anomalías en la configuración de los dilemas mismos, expuesta hace unas líneas, viciaría estos resultados.<sup>5</sup>

#### 4. REINTERPRETACIÓN DE LOS RESULTADOS Y DISTINCIÓN DE LOS DILEMAS

Para terminar este ensayo, quiero proponer un criterio diferente para distinguir los dilemas morales de los impersonales, así como una reinterpretación de los resultados obtenidos por los experimentos de Ciaramelli *et al.* y de Koenigs *et al.* Esto con el fin de diluir los problemas que he mostrado, para hablar a favor de la idea de que la CPV cumple un rol causal en la producción de juicios morales (al menos de algún tipo). Con ese objetivo en mente, resumiré brevemente los problemas señalados:

- a. En ambos experimentos hay un bajo nivel de aceptación por parte de los pacientes CPV de la acción propuesta en los dilemas morales personales. Si hubiese una anomalía en la aceptación de acciones utilitaristas, se esperaría que el rango de aceptación estuviera cercano al 100%, sin importar si los dilemas son de alto o bajo conflicto.
- b. En el experimento de Ciaramelli *et al.* hay una inconsistencia entre la idea de que la CPV anticipa emociones negativas y la alta aceptación por parte de sujetos normales en escenarios morales impersonales. Si es posible hacer una diferencia entre intención primaria y secundaria, tales escenarios supondrían una agencia de este último tipo, que provocaría la anticipación de emociones negativas.
- c. Dada la diferencia entre intención primaria y secundaria, el criterio propuesto por Green *et al.* (2001) para diferenciar dilemas morales de impersonales se hace insuficiente, pues los impersonales, al involucrar una agencia secundaria, podrían satisfacer todos los criterios de diferencia.

5 Un dato extraño, adicional, es que en los casos de dilemas morales impersonales, los sujetos normales no aprueban la acción propuesta como se esperaba que sucediera (cerca de un 89%). Como muestra la gráfica, apenas la aprobación alcanza 0.6. Esto se explica también por la configuración anómala de los dilemas (que también afecta los escenarios impersonales). Muchos de los dilemas no constituyen de hecho dilemas y se parecen al caso del hombre que no quiere manchar su carro. Esto también concuerda con el estudio de Kahane y Chackel (2008) según el cual sólo un 45% de los escenarios impersonales son dilemas genuinos.

- d. En el experimento de Koenigs *et al.*, no hay una diferencia estructural entre dilemas de alto y bajo conflicto, sino que la diferencia parece darse *ad hoc*. Por otra parte, algunos de los dilemas propuestos (tanto personales como impersonales) no revelan, de hecho, una estructura dilemática genuina.

Pues bien, la diferencia entre una y otra clase de escenarios morales, a mi juicio, debe entenderse de una forma más kantiana, propuesta que rechazan Green *et al.* (2001 2105-2106).

Tal división debería responder a que, en el caso de dilemas morales personales, se está usando a una persona como medio, mientras que en los dilemas morales impersonales no sucede esto, sino que el hecho de que se haga daño a otra persona es accidental. Esto no entra en conflicto con la diferencia propuesta entre agencia primaria y secundaria y tampoco se corresponde a ella. Una persona puede tener la intención de hacerle daño a una persona, sin que esto suponga que la está utilizando como medio para algo. También puede pasar que una persona sea responsable de manera secundaria de haberle hecho daño a alguien, sin que utilice a ese alguien como medio. Propongo entonces que, además de las tres condiciones de Green *et al.* para identificar violaciones morales personales, se tenga en cuenta la de utilizar a alguien como medio o como fin. El conjunto de condiciones sería: 1) hacer daño, 2) a una persona, 3) debe suponer una agencia (primaria o secundaria) y 4) utilizándola como medio. Esto, por supuesto, sigue la línea del análisis de Thomson (1985).

El postular este cuarto y último criterio para violaciones morales personales tiene como objetivo superar los problemas antes señalados en la interpretación de los resultados de los experimentos de Ciaramelli *et al.* y Koenigs *et al.* Tales problemas surgen teniendo como base la reunión de las tres primeras condiciones, pero no la última. El objetivo entonces es ver si añadiendo la condición kantiana, podemos superar los problemas de interpretación de los experimentos.

Por otro lado, recordemos que la predicción de Ciaramelli *et al.* y de Koenigs *et al.* supone que los pacientes CPV deben estar inclinados a la aprobación de la acción en los escenarios de dilemas morales personales. En otras palabras, los pacientes CPV deberían dar un juicio utilitarista en tales escenarios, pues al haber un daño en la CPV, las emociones no entrarían en juego, esto es, no provocarían una rápida respuesta para rechazar la acción propuesta. Tampoco entrarían en un conflicto con mecanismo cognitivos. De esta manera, el juicio producido por un paciente CPV, ante un dilema moral personal, sería utilitarista, ya que obedecería nada más que a un cálculo racional.

Pero la anormalidad en los juicios de pacientes CPV podría interpretarse, más bien, como teniendo un carácter ‘azaroso’. Con esto quizás sea posible concluir que, en efecto, hay un déficit en cuanto a la producción de juicios morales, pero que responde no a un cálculo puramente utilitarista, sino a una incapacidad, en sí misma, para expresar juicios morales (al menos en el caso de escenarios personales). Si esta lectura es posible, entonces vuelve a tomar fuerza la hipótesis de que la CPV y, por ende, las emociones cumplen un rol causal en la producción de juicios morales al menos de algún tipo particular. Propongo entonces, que la anormalidad sea leída en este sentido y no en el sentido utilitarista.

Ahora bien, en cuanto al primer problema, *i.e.*, el bajo rango de aprobación de respuesta por parte de los pacientes CPV en escenarios morales personales, al cambiar el sentido de anormalidad en la producción de juicio, la exigencia también cambia. Si hay una incapacidad, no se espera un rango anormal de respuesta utilitarista, en la respuesta a dilemas morales personales. Ya no debemos esperar un rango cercano al 100% en la aprobación de la violación propuesta en tales escenarios. En efecto, dado que la incapacidad que tendría un paciente CPV lo conduce a dar una suerte de respuestas ‘azarosas’, entonces el paciente CPV no tiene por qué presentar un rango de respuestas utilitaristas alto en ningún tipo de dilemas, sino que las respuestas deben presentarse más o menos uniformes (50-50). El rango de respuesta aprobatoria en ambos experimentos, en efecto, está muy cerca de 0.5.

En segundo lugar, no hay ningún problema con decir que la CPV anticipa emociones negativas sólo en casos de dilemas morales personales, pues esta restricción se podría explicar debido a que, de un modo por ejemplo intuitivo, el agente reconociera que no está usando a una persona como medio y, por tanto, la CPV no anticiparía emociones negativas en tal caso. Así, en el dilema clásico del tren, un sujeto sin problemas en la CPV, teniendo en cuenta los criterios de Green *et al.* y la diferencia propuesta entre intención primaria y secundaria, anticiparía emociones negativas y no aprobaría la acción. No obstante, el criterio propuesto de la utilización de alguien como medio, evitaría la activación de la emoción negativa. Lo anterior, asumiendo incluso que hay una agencia secundaria. Esto explicaría, ahora sí sin problemas, por qué los sujetos normales no se cohiben de aceptar la acción propuesta en los casos de dilemas morales impersonales.

En tercer lugar, evidentemente la diferencia entre juicios morales personales e impersonales ya no se hace oscura; más aún, es compatible con la diferencia propuesta entre agencia primaria y secundaria. Al introducir un cuarto criterio de distinción (utilizar o no a alguien como medio), la distinción entre ambos

dilemas es clara. La distinción entre agencia primaria y secundaria no es, como se mostró, análoga a la distinción entre usar a alguien como medio o no.

Por último, dado que los resultados de tales experimentos podrían leerse de ese modo ‘azaroso’, se elimina el problema de la división *ad hoc* entre dilemas de alto y bajo conflicto. No hay problema con el hecho de que los pacientes CPV presenten una respuesta normal ante dilemas de la última clase y no ante dilemas de la primera. Eso podría ser producto de la incapacidad misma del paciente para juzgar (lo cual dejaría de lado la hipótesis de la anormalidad utilitarista en las respuestas del paciente CPV). Esto, en conjunción con la mala configuración de los dilemas utilizados por Koenigs *et al.* que se expuso con anterioridad, explicaría entonces las respuestas dadas en ese experimento. Sin embargo, pienso que el experimento de Koenigs *et al.*, al presentar problemas de este tipo, no puede ser adecuadamente reinterpretado. Habría que armar una batería con una configuración clara (teniendo en cuenta los criterios establecidos) y, con tal batería, realizar un experimento similar para contrastar los resultados.

## 5. CONCLUSIONES

He analizado varios experimentos que intentaban mostrar el rol causal de la CPV —y por ende— de las emociones en el juicio moral. Green *et al.* (2001, 2004), muestran que hay activación de la CPV cuando los sujetos contemplan dilemas morales personales (y la activación es más notoria en los dilemas difíciles). Sin embargo, la activación podría ser accidental y no esencial (causal) al juicio moral. Los experimentos de Ciaramelli *et al.* y de Koenigs *et al.*, muestran que debe haber, en efecto, un rol causal de la CPV ante escenarios morales personales, pues pacientes CPV muestran una anormalidad en la producción de juicios en tales escenarios. Mostré algunos problemas de tales experimentos. Por un lado, la respuesta no fue tan anormal como lo esperado. En segundo lugar, la anticipación de acciones negativas por parte de la CPV no era consistente con el hecho de que sujetos normales aprobaran juicios morales impersonales (ya que estos supondrían también una agencia). Además, la diferencia entre juicios morales personales e impersonales parecía oscura. Por último, en uno de los experimentos, los dilemas usados no presentaban una configuración adecuada, lo cual viciaría los resultados. Propuse entender la diferencia de los dilemas morales personales y los impersonales, a partir de un criterio diferente (la utilización o no de una persona como medio). Asimismo, propuse una reinterpretación de lo que debe entenderse por respuesta anormal en el caso de los pacientes CPV ante escenarios morales

personales. Tal anormalidad no sería la muestra de un patrón utilitarista, sino de una incapacidad para producir juicios. Estas dos propuestas solucionarían los problemas señalados, aunque el caso de la configuración problemática de dilemas queda aún abierto. La hipótesis según la cual la CPV juega un rol causal en el juicio moral, no obstante, es favorecida.

## TRABAJOS CITADOS

- Ciaramelli *et al.* "Selective Deficit in Personal Moral Judgment Following Damage to Ventromedial Prefrontal Cortex". *Social Cognitive and Affect Neuroscience*. 2. (2007): 84-92.
- Bechara, Antoine & Damasio, Antonio. "The somatic Marker Hypothesis: A Neural Theory of Economic Decision". *Games and Economic Behavior*, 52. (2005): 336-372.
- Green *et al.* "An fMRI Investigation of Emotional Engagement in Moral Judgment". *Science*, 293 (2001): 2105-2108.
- Green *et al.* "The Neural Bases of Cognitive Conflict and Control in Moral Judgment". *Neuron*, 44 (2004): 389-400.
- Kahane, Guy & Shackel, Nicholas. "Do Abnormal Responses Show Utilitarian Bias?" *Nature*, 452. 2008.
- Koenigs *et al.* "Damage to the Prefrontal cortex Increases Utilitarian Moral Judgments". *Nature Publishing Group*, 446 (2007): 908-911.
- Rosas, Alejandro; Caviedes, Esteban; Arciniegas, Alejandra & Arciniegas, Andrea. "¿Decisión utilitarista o decisión aleatoria? Crítica a una tesis atrincherada en la neurociencia cognitiva". *Ideas y Valores* 62 (2013):179-199.
- Thomson, Judith. "The Trolley Problem". *The Yale Law Journal*, 94 (1985): 1395-1415.
- Young *et al.* "Damage to Ventromedial Prefrontal Cortex impairs Judgment of Harmful Intent". *Neuron*, 10162 (2010): 1-7.

