TESIS DOCTORAL

EFECTOS DEL PROCESAMIENTO EMOCIONAL SOBRE LA IMPULSIVIDAD EN LA TOMA DE DECISIONES

David Contreras Ros

Dirigida por

Andrés Catena Martínez

Antonio Maldonado López

Programa de doctorado de Psicología Experimental y Neurociencias del Comportamiento

CIMCYC

Universidad de Granada

Editor: Editorial de la Universidad de Granada
Autor: David Contreras Ros
D.L.: GR 378-2014
ISBN: 978-84-9028-775-0

ÍNDICE

Prefacio

Impulsividad y control cognitivo

Como tantos otros términos empleados en psicología, la palabra "impulsividad" es un término prestado del lenguaje común, lo que hace que acarree múltiples connotaciones y no exista una definición operacional única del mismo. Cualquier persona es capaz de dar ejemplos intuitivos de lo que significa "ser impulsivo" o realizar un "acto impulsivo": beber una copa más, hacer una llamada telefónica a deshora, improvisar un viaje de un día para otro. La definición de lo que es o no es impulsivo depende de la cultura y ha variado a lo largo de la Historia, lo cual no es un inconveniente para el uso informal del término, pero supone un problema más grave para su uso científico (Evenden, 1999). Sin embargo, y a pesar de que hay considerable ambigüedad respecto a lo que significa impulsividad en el vocabulario científico, el término se ha empleado de forma profusa tanto para describir los síntomas como para proponer explicaciones causales de numerosos trastornos. Por ejemplo, existen varias teorías que consideran el trastorno de déficit de atención/hiperactividad como la consecuencia de una incapacidad para la "inhibición ejecutiva" de respuestas dominantes (Nigg, 2000). Así mismo, las adicciones se han considerado como la consecuencia de la incapacidad de la persona para sucumbir a la tentación del placer inmediato, a pesar de las consecuencias negativas a largo plazo que tiene el consumo de la sustancia adictiva, un patrón de conducta que también se ha descrito como impulsivo (Olmstead, 2006; Perry & Carroll, 2008). La conducta agresiva (McKay & Halperin, 2001) y los trastornos de la conducta alimentaria (Carrard, Crépin, Ceschi, Golay, & Van der Linden, 2012; Dawe & Loxton, 2004) también se han relacionado con la dificultad para "controlar los impulsos".

Una definición de manual del término impulsividad afirma que "abarca acciones que parecen concebidas de manera pobre, expresadas prematuramente, excesivamente arriesgadas o inapropiadas para la situación, y que, a menudo, resultan en consecuencias indeseables" (Daruna & Barnes, 1993, p. 23). La definición nos parece interesante porque avanza la idea de que la impulsividad no es un concepto monolítico, sino una variedad de aspectos diferenciables del comportamiento. En la definición aparecen, de forma implícita, la falta de planificación ("concebidas de manera pobre"), la tendencia a llevar a cabo las acciones antes de disponer de la suficiente información para tomar la decisión adecuada, lo que también podría llamarse urgencia ("expresadas prematuramente"), la ejecución de acciones de forma automática, estereotipada o inflexible, sin tener en consideración el contexto ("inapropiadas para la situación") y la propensión al riesgo ("excesivamente arriesgadas").

Todos estos tipos de impulsividad merecen consideración separada, ya que existe evidencia de que su sustrato neurofisiológico es diferente (Evenden, 1999; Winstanley, Theobald, Dalley, Cardinal, & Robbins, 2006, Dalley et al., 2011).

Mientras que el término impulsividad es de uso más común en el ámbito de la psicología clínica, en el campo de la psicología cognitiva es más habitual oír hablar de "control cognitivo" o "control ejecutivo". Impulsividad es un sinónimo aproximado de control cognitivo deficient (Dalley, Everitt, & Robbins, 2011; Grant & Potenza, 2011). El abanico de conductas que se consideran impulsivas coincide en buena medida con las que regulan los mecanismos de control cognitivo (de Wit, 2009; Dick et al., 2010; Stautz & Cooper, 2013; Thamotharan, Lange, Zale, Huffhines, & Fields, 2012). Este concepto, el de control cognitivo tiene que ver con la capacidad de los organismos superiores para llevar a cabo patrones de conducta flexibles, en los que la respuesta escogida no está rígidamente determinada por la identidad del estímulo sensorial percibido, sino que se elige conforme al contexto y a las metas a largo plazo del organismo (Caswell, Morgan, & Duka, 2013; Miller & Cohen, 2001; Miller, 2000). Es un concepto, por tanto, relacionado con la capacidad para planificar, para adecuar las acciones al contexto, y para representar las metas a largo plazo, incluyendo la capacidad para representar el *valor* de la meta.

Tipos de impulsividad

En este trabajo diferenciamos cuatro características de la conducta que pueden etiquetarse de "impulsivas" (Kim & Lee, 2011; Robbins, Gillan, Smith, de Wit, & Ersche, 2012). Algunos de las siguientes descripciones se elaboran más en detalle en el capítulo de introducción:

- Incapacidad para soportar la demora de un reforzador: aunque todos los organismos prefieren un reforzador administrado de forma inmediata frente a uno demorado cuando ambos tienen el mismo valor, lo que implica que aplican un descuento temporal a las recompensas futuras (Critchfield & Kollins, 2001), en general se etiqueta como más impulsiva aquella conducta que aplica un descuento temporal mayor, es decir, la que prefiere reforzadores inmediatos de menor cuantía frente a uno demorado de cuantía fija. Ballard & Knutson, (2009) demuestran la implicación de dos subsitemas, ligado uno a las proyecciones dopaminégicas mesolímbicas (núcleo accumbens) hacia la corteza preforntal mesial (MPFC) y cingulado posterior (PCC), cuya activación está relacionada con la magnitud de la demora futura, y otro que incluye al dorsolateral prefrontal (DLPFC) y parietal posterior (PPC) cuya actividad está negativamente

relacionada con esa misma variable.

- Incapacidad para cancelar respuestas erróneas ya en curso (impulsividad motora). En situaciones en las que una respuesta ya se ha iniciado pero debe detenerse por un cambio en las circunstancias, o porque el organismo adquiere nueva información, detener la respuesta en curso a tiempo es crucial para la supervivencia. La dificultad para lograrlo se puede considerar un tipo de impulsividad (Aron, 2007; Caswell et al., 2013).

- Tendencia a ejecutar respuestas guiadas por la presencia e identidad o saliencia del estímulo, sin atender al contexto o a las metas que el organismo persigue (Christiansen, Cole, Goudie, & Field, 2012), es decir, un fallo en el control *top-down* o cognitivo de la conducta (Miller & Cohen, 2001). Es la tendencia que se observa en el trastorno neuropsicológico de "conducta de utilización" (Lhermitte, 1983).

- Tendencia a ejecutar respuestas antes de haber adquirido información suficiente para decidir cuál es la adecuada (urgencia o impulsividad en la preparación). En todas las decisiones existe un balance entre la conveniencia de emplear poco tiempo en decidir y la de poseer información suficiente para elegir el mejor curso de acción. Un umbral de decisión bajo llevará a tomar decisiones más rápidas, pero con más errores. Un umbral más elevado exigirá más certeza para tomar la decisión, pero también conllevará más tiempo. Se ha comprobado que los humanos tienden a ajustar el umbral de decisión de forma que la tasa de reforzamiento global, teniendo en cuenta los errores ocasionales, sea máxima (Simen et al., 2009). Se puede considerar más impulsiva una conducta cuyo umbral de decisión es más bajo (Kim & Lee, 2011).

Influencias de la emoción sobre la impulsividad

El peso que se atribuye a la impulsividad en la etiología de numerosos trastornos de conducta ha promovido numerosas investigaciones sobre las bases neurofisiológicas y genéticas de la impulsividad, que persiguen identificar poblaciones de riesgo y el desarrollo de fármacos adecuados para su control. La perspectiva que adoptan estas investigaciones es la de que la impulsividad es un rasgo de personalidad del individuo. Algunos de estos estudios atribuyen específicamente la impulsividad a la incapacidad para regular las emociones. Por ejemplo, Davidson, Putnam, & Larson (2000) afirman que la violencia impulsiva se debe a una disfunción en el circuito neuronal que regula las emociones. El hecho de que se atribuyan conductas impulsivas a un fallo en la regulación de las emociones, no

obstante, supone admitir implícitamente que las emociones puntuales también influyen en la impulsividad de la conducta. Es decir, si bien la propensión a la conducta impulsiva puede ser un rasgo de personalidad, también es cierto que las emociones puntuales pueden provocar conductas impulsivas de forma transitoria.

Existen datos empíricos que llevan a pensar que así es, en efecto. McClure, Laibson, Loewenstein, & Cohen, 2004 (véase también Eppinger, Nystrom, & Cohen, 2012) midieron mediante resonancia magnética funcional la actividad cerebral de personas mientras escogían entre refuerzos monetarios inmediatos de pequeño importe y demorados de mayor importe, y observaron que la valoración de los refuerzos inmediatos implicaba áreas emocionales del cerebro (sistema límbico), mientras que la estimación de los demorados activaba áreas cognitivas (prefrontal dorsolateral, corteza parietal posterior). Similares patrones de activación se observaron también con reforzadores primarios (zumo y agua) (McClure, Ericson, Laibson, Loewenstein, & Cohen, 2007). Especialmente interesante es el hecho de que preactivar las áreas emocionales (permitiendo la visión del objeto ofrecido posteriormente como reforzador, o presentándoles su olor) provoca un sesgo hacia las decisiones impulsivas (McClure et al., 2004; Metcalfe & Mischel, 1999), lo que indica que las emociones agudas influyen directamente en la conducta impulsiva.

También hay razones teóricas para pensar que las emociones interfieren con la urgencia con la que se toman decisiones. Es un hecho conocido que un animal al que se le presenta un estímulo condicionado asociado con la aparición de bebida pulsa con más frecuencia una palanca que le proporciona comida (Dickinson & Dawson, 1987). Una explicación de este fenómeno es que el estímulo condicionado provoca que el animal espere la aparición del reforzador asociado, y eso crea en el animal un sentido de urgencia que incrementa de forma indiscriminada la frecuencia y la rapidez con la que lleva a cabo conductas. Pues bien, se han presentado argumentos teóricos que apoyan la idea de que la existencia en el entorno de una oportunidad para obtener un reforzador o la presencia de una amenaza hace que, de hecho, sea óptimo llevar a cabo *cualquier* acción más deprisa (Niv, Daw, Joel, & Dayan, 2007), ya que cada segundo desperdiciado implica un coste de oportunidad mayor que en ausencia de esa oportunidad o amenaza.

Otras influencias de la emoción sobre la conducta

En estudios conductuales de los efectos de la emoción, un factor de confusión es el hecho de que la emoción puede influir sobre la conducta por diversas vías, de las cuales su posible efecto en la

impulsividad es solamente una. Existe considerable evidencia empírica de que las emociones, al margen de su posible influencia sobre la impulsividad, influyen directamente sobre aspectos de bajo nivel de la percepción sensorial (Padmala & Pessoa, 2008, 2011), incrementando la sensibilidad hacia ciertos estímulos y reduciéndola hacia otros (por ejemplo, mejorando la sensibilidad a enrejados de baja frecuencia espacial espacial y empeorándola para los de alta (Bocanegra & Zeelenberg, 2009), lo que en último término influye en los tiempos de reacción y las tasas de error. De la misma forma, se ha comprobado que los estímulos emocionales tienden a capturar la atención (Vuilleumier & Huang, 2009; Vuilleumier, 2005), por lo que es menor el tiempo necesario para detectarlos entre distractores (Ohman, Flykt, & Esteves, 2001), incrementan el tiempo de reacción cuando son irrelevantes para la tarea al mantener la atención enganchada más tiempo que otros distractores neutros (Fox, Russo, & Dutton, 2002), o provocan periodos refractarios para la detección de estímulos que aparecen poco después (lo que se conoce como parpadeo atencional emocional (Most, Chun, Widders, & Zald, 2005).

Las emociones preactivan también programas motores específicos. Ciertos teóricos prominentes de la emoción, las conciben, de hecho, como estados de disposición para la acción (Frijda, 1986). Numerosos resultados científicos apoyan la noción de que las emociones interfieren con la propensión del organismo a llevar a cabo determinadas acciones. Los estudios de etología han documentado que el miedo provoca en los mamíferos patrones de conducta característicos: una paralización inicial (*freezing*) seguida de lucha o huida, dependiendo de la distancia al peligro y de la existencia o no de vías de escape (Blanchard & Blanchard, 1988). La paralización provocada por el miedo se ha estudiado también ampliamente en laboratorio usando el paradigma de condicionamiento del miedo, y el circuito neuronal del que depende esta paralización se conoce con cierto detalle (LeDoux, 2000; Phelps & LeDoux, 2005). Por otra parte, el valor afectivo positivo o negativo de los estímulos facilita respectivamente respuestas de aproximación o retirada, como se ha observado en experimentos en los que los participantes debían efectuar respuestas de retirada (e.g., tirar de una palanca) o de aproximación (e.g., empujarla) al detectar la aparición de un estímulo visual, y las aproximaciones eran más rápidas para estímulos agradables que para desagradables y viceversa (Chen & Bargh, 1999; Duckworth, Bargh, Garcia, & Chaiken, 2002).

Por último, también se ha considerado que la emoción y la motivación pueden afectar la actividad cerebral incrementando la intercomunicación entre regiones corticales y subcorticales, preferentemente (Kinnison, Padmala, Choi, & Pessoa, 2012). Es importante, en consecuencia, tener en mente que son múltiples las interpretaciones posibles de los efectos de la emoción sobre las variables medidas en

experimentos conductuales (tiempos de reacción, tasas de errores).

Algunas tareas empleadas para medir la impulsividad

Las tareas experimentales más empleadas para estudiar la impulsividad motora son la tarea *go*/*no-go* (empleada ya por (Donders, 1868/1969) y la tarea *stop-signal* o tarea de señal de parada (Logan & Cowan, 1984; Logan, 1985). En la tarea *go*/*no-go* el participante debe efectuar una respuesta motora (pulsar un botón, o tocar una pantalla) cuando se le muestra cierto estímulo (ensayos *go*), y abstenerse de llevar a cabo dicha respuesta cuando aparece otro estímulo diferente (ensayos *no*-go). Para crear en el participante una propensión a responder en todos los ensayos, el número de ensayos *go* normalmente es muy superior al de ensayos *no-go*. De ese modo, en los ensayos no-go, que son infrecuentes, el participante debe activamente refrenar el impulso de responder. La variable dependiente conductual que se estudia habitualmente es la tasa de errores cometidos en ensayos *no-go*, que se interpretan como fallos en la inhibición de respuestas motoras.

Esta interpretación presupone que la ausencia de respuestas en los ensayos no-go se debe a la actuación de un mecanismo cerebral que activamente detiene una respuesta que, de otro modo, se habría producido, y no a la mera ausencia de esta respuesta. Este supuesto ha obtenido cierto apoyo empírico en experimentos de neuroimagen que han mostrado que durante los ensayos no-go se activan ciertas áreas cerebrales, inactivas durante los ensayos go, que forman parte de una red que se ha relacionado con la inhibición de respuestas motoras (Aron, 2007).

En la tarea *stop-signal* se emplean estímulos diferentes para requerir la acción del participante y para indicarle que la detenga. La señal de parada (la que le indica que no lleve a cabo la acción) se presenta demorada respecto del target, con una demora variable, y en ocasiones, en una modalidad sensorial diferente (por ejemplo, el target puede ser un estímulo presentado visualmente en una pantalla, y la señal de parada, un sonido). A mayor demora entre target y señal de parada, lógicamente, más difícil resulta detener la respuesta que se inicia ante el target. Manipulando esta demora puede estimarse hasta qué momento es capaz la persona de interrupir la acción ya lanzada, y eso da una medida indirecta de su capacidad para inhibir respuestas.

Métodos para manipular la emoción

Una distinción frecuente en la literatura científica sobre las emociones es la que se hace entre estados

de ánimo (*moods*) y emociones puntuales (*emotions*) (Gross, 1998). Mientras que las emociones puntuales son fenómenos agudos, normalmente dirigidos a un objeto concreto, acompañados de una intensa respuesta fisiológica y tendencias de acción inmediata, los estados de ánimo abarcan periodos más largos de tiempo, son más difusos, e influyen más sobre la cognición que sobre las acciones. Nuestro interés en este trabajo se centra en el efecto inmediato de las emociones puntuales, más que en el de los estados de ánimo.

Se han empleado múltiples estrategias para la inducción de estados emocionales. Desde el pionero estudio de Landis (1924) en el que empleó técnicas tan variopintas como pedir al participante que metiera la mano en un cubo con ranas o que contemplara un cuadro de un crucificado, la variedad de herramientas disponible para provocar emociones no ha parado de aumentar. En experimentos de psicología social se ha empleado el insulto como forma de inducir ira y la comparación social y el elogio como forma de inducir alegría o tristeza (Harmon-Jones, Amodio, & Zinner, 2007). Para inducir asco o diversión se han empleado fragmentos de películas (Gross & Muñoz, 1995; Gross & Thompson, 2007). La lectura de narraciones, la presentación de reforzadores primarios (sabores, olores, dolor, o la amenaza del dolor), escuchar piezas musicales (Zentner, Grandjean, & Scherer, 2008), la hipnosis (Whorwell, Houghton, Taylor, & Maxton, 1992) e instruir al participante para que configure ciertas expresiones faciales (Adelmann & Zajonc, 1989), son otros métodos de inducción emocional que se han utilizado en investigación. Estos métodos difieren en el grado de control que proporcionan al experimentador sobre el momento en el que la emoción se induce, y sobre su duración. Algunos, por su propia naturaleza, requieren tiempos más largos para aplicarlos, y provocan efectos más extendidos en el tiempo (más parecidos a un estado de ánimo que a una emoción), mientras que otros son de efecto instantáneo y más breve.

Para nuestras manipulaciones emocionales nos servimos de fotografías y sonidos emocionales de dos catálogos cuya capacidad para inducir emociones se ha baremado en diversas poblaciones, el sistema internacional de imágenes afectivas (IAPS; Lang, Bradley, & Cuthbert, 1999; Moltó el al., 1999) y el sistema internacional de sonidos afectivos (IADS; Bradley & Lang, 1999). Su uso tiene la ventaja de que permiten un control muy preciso del instante en el que comienza el procesamiento emocional, y tienen un efecto de duración limitada.

El componente N2 como Indicador electrofisiológico del control cognitivo

En los últimos 30 años, numerosos estudios con potenciales evocados han documentado la existencia de un componente negativo mediofrontal relacionado con el control cognitivo (el "N2 cognitivo"), cuyo máximo de actividad tiene lugar unos 250 ms después del onset del estímulo, y cuyo origen probable se encuentra en la corteza prefrontal medial. Este componente aparece cuando el participante tiene que vencer una propensión a responder para llevar a cabo una tarea experimental correctamente. Por ejemplo, en la tarea *stop-signal*, la señal de parada elicita la aparición de un N2 que no aparece en los ensayos sin parada, y este N2 aparece más tarde en los ensayos en los que el participante es incapaz de detener la respuesta (Kok, Ramautar, De Ruiter, Band, & Ridderinkhof, 2004). En tareas *go/no-go*, se observa un componente N2 más amplio en los ensayos *no-go* que en los ensayos *go*, incluso aunque ninguno de ellos requiera una respuesta muscular, lo que descarta que el efecto se deba a la ausencia de respuesta muscular en los ensayos *no-go* (Pfefferbaum, Ford, Weller, & Kopell, 1985). En la tarea de flancos de Eriksen, (1995), donde el participante debe determinar la identidad de un estímulo flanqueado por otros que pueden ser idénticos (condición congruente) o diferentes (condición incongruente), se genera un N2 en los ensayos incongruentes más amplio que en los congruentes, presumiblemente porque el participante tiene que vencer en ocasiones la tendencia a responder con la mano correspondiente a los estímulos que flanquean al target (Coles, Gratton, Bashore, Eriksen, & Donchin, 1985; Gehring, Gratton, Coles, & Donchin, 1992). De hecho, la amplitud del N2 crece con el número de estímulos incongruentes que flanquean al target (Forster, Carter, Cohen, & Cho, 2011).

Aunque la mera frecuencia de los estímulos parece suficiente para provocar un N2, dado que los estímulos infrecuentes también lo evocan en los paradigmas oddball, en los que un target infrecuente se presenta intercalado en un tren de distractores frecuentes (Folstein & Van Petten, 2008), el N2 visible en las tareas de control cognitivo parece tener un origen distinto, ya que el número de ensayos congruentes e incongruentes es el mismo en la tarea de flancos, y en la tarea go/no-go, el N2 es visible incluso cuando los ensayos go y no-go son equiprobables (Kok, 1986, p. 198; Nieuwenhuis, Yeung, Van Den Wildenberg, & Ridderinkhof, 2003).

Aunque existe consenso sobre el hecho de que este componente tiene relación con el control cognitivo, existen varias teorías contrapuestas sobre su significado funcional, que lo han relacionado con la detección de conflictos entre respuestas (Botvinick, Cohen, & Carter, 2004), con la inhibición de respuestas motoras en curso (Jodo & Kayama, 1992; Pfefferbaum et al., 1985) o con la detección de

errores y la evaluación de los propios actos (Alexander & Brown, 2011; Brown & Braver, 2005). Paradójicamente, todas estas teorías cuentan con considerable apoyo empírico, aunque ninguna de forma concluyente.

Objetivos de la tesis

Esta tesis doctoral pretende determinar si el procesamiento de imágenes y sonidos emocionales tiene efectos inmediatos sobre la impulsividad y la conducta de riesgo. El propósito es determinar la cronometría precisa de estos efectos, empleando tanto tareas controladas de laboratorio como otras de mayor validez ecológica, que ilustren las consecuencias de estos efectos impulsivos en la vida real.

Dado que en los estudios conductuales sólo se observa el resultado final de los procesos cognitivos (esto es, la propia conducta), y resulta difícil discernir qué procesos han intervenido para alterarla, la segunda parte de la tesis indaga sobre cuáles son las posibilidades de la electroencefalografía para obtener medidas más directas del control cognitivo. En primer lugar se intentará responder a la pregunta sobre si el N2 puede intervenir en la cancelación *top-down* de respuestas erróneas ya iniciadas. En segundo lugar, se presenta un nuevo método para la generación de filtros espaciales a medida para aislar procesos cerebrales específicos.

Resumen

El capítulo 1, introductorio, revisa los conceptos de impulsividad y propensión al riesgo. En él se hace una revisión de la literatura sobre las posibles funciones de una de las áreas cerebrales a las que se atribuye el control de la interacción entre emoción e impulsividad, y la responsabilidad de integrar y representar el valor hedónico de las consecuencias esperadas de las acciones.

En el capítulo 2 se estudian los efectos inmediatos del procesamiento imágenes emocionales sobre la capacidad para inhibir respuestas motoras, mediante una tarea *go/no-go*. En él se informa de la existencia de un efecto de facilitación del procesamiento emocional sobre la generación de respuestas que no se había descrito previamente, y que puede interpretarse como un aumento transitorio en la impulsividad de preparación provocado por la emoción.

El capítulo 3 investiga los efectos del procesamiento de sonidos emocionales en el comportamiento de conducción empleando un simulador de motocicleta, y se analizan las consecuencias de la presentación de estos sonidos sobre la velocidad del conductor y la probabilidad de sufrir un accidente en situaciones

de riesgo. Dado que las tareas típicas, mencionadas más arriba, pueden considerarse de escasa validez ecológica, el valor añadido de este trabajo fue el uso de simuladores de media fidelidad, que permiten acercarse al estudio del procesamiento emocional en situaciones más cercanas a las reales.

El capitulo 4 es el resultado final de varios estudios exploratorios en los que se partió de la idea de que el proceso responsable de la supresión o inhibición de una respuesta motora debe actuar, siguiendo un orden causal, antes de que dicha respuesta se haya cancelado. Utilizando tres tareas típicas en las que está implicada la supresión motora (*go/no-go*, tarea de flancos de Eriksen y *stop-signal*) se midió la actividad muscular durante la supresión de respuestas erróneas (amagos de respuesta) simultáneamente a la actividad EEG. Tanto a escala grupal, como a la del participante individual, como ensayo a ensayo, se observó que el inicio del componente N2 es posterior al comienzo de la cancelación del amago de respuesta. Además, el área acumulada bajo la curva EMG previamente al inicio del componente N2 fue significativamente mayor del 50%. Estos resultados permiten concluir que el N2 (y por tanto, cualquier otro componente cognitivo posterior) no puede considerarse un índice del proceso encargado de la supresión/inhibición de la respuesta parcialmente ejecutada, aunque no descarta que puede reflejar el procesamiento responsable del ajuste de la estrategia de respuestas para los ensayos subsiguientes.

Finalmente en el capítulo 5 se presenta una posible solución a uno de los problemas más acuciantes para la interpretación del significado de los componentes en la onda ERP: el hecho de que las actividades electrofisiológicas generadas por diversos procesos mentales (perceptivos, motores, etc.) están superpuestas en la onda registrada. Nuestro objetivo comenzó siendo la obtención de una medida más limpia de los componentes motores –para medir el efecto sobre ellos de la inhibición– y del componente N2. Nuestra aportación parte de supuestos que son comúnmente aceptados en la literatura psicológica, particularmente la linealidad del medio cerebral y de los procesos que dan lugar a la onda observada. Estos mismos supuestos subyacen a otras técnicas de separación de fuentes, como el análisis de componentes independientes (ICA). El ICA es quizá el método de separación que tiene más aceptación en la actualidad. Sin embargo, aunque muy eficaz para aislar fuentes no cerebrales (ruido procedente del suministro eléctrico o de movimientos oculares), en nuestros análisis el ICA ha demostrado de forma reiterada de ser incapaz de separar componentes como el N2 y el P3, debido a que sus amplitudes suelen covariar, a pesar de que estos componentes son disociables anatómica y funcionalmente. Además, el ICA y, en general, todos los métodos conocidos como de separación ciega de fuentes (*blind source separation*), llevan a cabo sus análisis sin servirse de ninguna información conocida a priori acerca de los procesos que pueden estar implicados en la producción de la onda EEG.

Si bien el enfoque de la separación ciega de fuentes tiene el atractivo de ser teóricamente neutral, implica despreciar todo el conocimiento psicológico y neurológico previo sobre los procesos objeto de investigación, que podría usarse para realizar conjeturas sobre el curso temporal de actividad de las fuentes que podrían intervenir en una tarea concreta. Nuestro método trata de aprovechar ese conocimiento para realizar una separación guiada de los componentes mezclados en la onda EEG, y en el capítulo 5 demostramos su capacidad para aislar la actividad cerebral relacionada con la respuestas motoras ejecutadas por cada mano por separado, así como la de dos de los llamados componentes "cognitivos", concretamente el N2 y P3.

En conclusión, la presente tesis doctoral es el resultado de una trayectoria investigadora que abarca muchos años de exploración en campos imbricados en los que consideramos que se realizan algunas aportaciones fundamentales de naturaleza teórica y metodológica. Entre estas aportaciones destacan el efecto de las manipulaciones emocionales sobre tareas típicas de laboratorio (*go/no-go*) y tareas ecológicas (conducción de simuladores de moto). Este efecto emocional consiste básicamente en un incremento en la urgencia para responder, que lleva a la facilitación en la tarea *go/no-go* y a un incremento de accidentes en la tarea del simulador. Además, hemos podido demostrar que un índice putativo de procesos de supresión/inhibición como el componente N2 no parece estar implicado en la supresión ensayo a ensayo de respuestas inadecuadas o erróneas (descartando, de paso, la implicación de cualquier otro componente posterior en esa misma función). Finalmente, nuestra aportación metodológica ha consistido en la presentación de un método de separación guiada de componentes en la onda ERP que podría contribuir a mejorar la forma en que se analizan los resultados de los estudios con EEG.

Summary in English

Chapter 1 is an introductory chapter that reviews the concepts of impulsivity and risk preference. The chapter reviews the scientific literature about the possible functions of the area of the brain that is believed to control the interactions between emotion and impulsivity, and the representation and integration of the hedonic value of expected consequences of actions.

Chapter 2 studies the short-term effects of the processing of emotional images on the ability to inhibit motor responses, using a go/no-go task. The chapter reports the existence of a facilitation effect of emotional processing on response execution, an effect that had not been described before, and that can be interpreted as a transient increment in preparation impulsivity caused by emotion.

Chapter 3 investigates the effects of processing emotional sounds on driving behavior using a motorcycle simulator, and the consequences of such processing on the speed chosen by the rider and on the probability of suffering an accident in a risk scenario are analyzed. Since typical lab tasks, described above, can be considered of low ecological validity, the added value of this work was the use of medium-fidelity simulator, that enable to approach the study of emotional processing in situations closer to real life ones.

Chapter 4 is the outcome of several exploratory studies that started from the idea that the process responsible for the suppression or inhibition of a motor response must act, following causal order, *before* that response is canceled. Using three classical experimental tasks that require response cancellation (go/no-go task, Eriksen flankers task and stop-signal task), muscular activity was measured during incorrect response suppression (during partial responses), simultaneously with EEG activity. At the group level, as well as at the individual subject level, and at the trial-by-trial level, it was observed that the onset of the N2 component lags the beginning of the cancellation of the associated partial response. Moreover, the area under the EMG curve and before the N2 onset was significantly larger than 50% of the total area under the EMG curve for that response. Those results lead us to conclude that the N2 (and hence any other posterior cognitive component) cannot be considered an index of the process responsible for the suppression/inhibition of partial responses, although this doesn't rule out the possibility that it reflects the mental activity responsible for the adjustment of the level of cognitive control exerted in subsequent trials.

Finally, chapter 5 introduces a possible solution to one of the most pressing problems for the interpretation of the meaning of components in ERP waveforms: the fact that the electrophysiological activities evoked by different mental processes (perceptive, motor, etc.) overlap in the recorded waveform. Our original aim was to obtain a cleaner measure of motor components –so as to evaluate the effect of inhibition on them–, and of the N2 component. Our contribution is based on assumptions commonly accepted in the psychological literature, especially on the linearity of the brain milieu and of the electrical processes that give rise to the observed waveform. These same assumptions underlie other source separation techniques, such as independent component analysis (ICA). ICA is perhaps the most accepted source separation method these days. However, although very effective to isolate non-brain electrical sources (such as mains noise or eye-movement related noise), in our analyses ICA has repeatedly proved unable to separate components such as the N2 and the P3, because their amplitudes tend to covary, despite those components being anatomically and functionally dissociable. Besides,

ICA, and all blind source separation methods generally, perform their analyses without making use of any previously known information about the processes that are likely involved in the generation of the EEG waveforms. Although the approach of blind source separation methods has the appeal of being theoretically neutral, it implies neglecting all previous psychological and neurological knowledge about the processes under study, a knowledge that could be used to make conjectures about the temporal course of activity of the brain sources that might take part in a certain task. Our method tries to leverage on that knowledge to make a guided separation of the components mixed in the EEG waveform, and in chapter 5 we prove its ability to isolate the brain activity associated with motor actions of each hand separated, as well as that of two cognitive components, namely the N2 and the P3.

In sum, the present doctoral thesis is the result of a research trajectory that has extended along several years of exploration of interrelated fields, to which we believe sound contributions, both theoretical and methodological, have been made. Among these contributions we may highlight the study of the effect of emotional manipulations on classical experimental tasks (go/no-go, stop-signal task) and ecological tasks (driving a motorcycle simulator). The said emotional effect is basically an increase in the urgency to respond, that leads to facilitation in the go/no-go task, and to an increase of the rate of accidents in the simulator task. Moreover, we have been able to show that a purported index of suppression/inhibition processes –the N2 component– cannot be involved in the intra-trial cancellation of incorrect or inappropriate responses (ruling out at the same time the involvement of any other posterior component in the same function). Finally, our methodological contribution has been the introduction of a method for the guided separation of components in the ERP waveform that could contribute to improve the way in which data from EEG studies are analyzed.

Referencias

Adelmann, P. K., & Zajonc, R. B. (1989). Facial efference and the experience of emotion. *Annual review of Psychology*, *40*(1), 249-280.

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature neuroscience*, *14*(10), 1338-1344.

Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *Neuroscientist*, *13*(3), 214-228.

Ballard, K., & Knutson, B. (2009). Dissociable neural representations of future reward magnitude and delay during temporal discounting. *Neuroimage*, *45*(1), 143.

Blanchard, D. C., & Blanchard, R. J. (1988). Ethoexperimental approaches to the biology of emotion. *Annual review of psychology*, *39*(1), 43-68.

Bocanegra, B. R., & Zeelenberg, R. (2009). Emotion improves and impairs early vision. *Psychological Science*, *20*(6), 707-713.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in cognitive sciences*, *8*(12), 539-546.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*(5712), 1118-1121.

Carrard, I., Crépin, C., Ceschi, G., Golay, A., & Van der Linden, M. (2012). Relations between pure dietary and dietary-negative affect subtypes and impulsivity and reinforcement sensitivity in binge eating individuals. *Eating behaviors*, *13*(1), 13-19.

Caswell, A. J., Morgan, M. J., & Duka, T. (2013). Inhibitory Control Contributes to «Motor»-but not «Cognitive»-Impulsivity. *Experimental Psychology (formerly Zeitschrift für Experimentelle*

Psychologie), 1-11.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, *25*(2), 215-224.

Christiansen, P., Cole, J. C., Goudie, A. J., & Field, M. (2012). Components of behavioural impulsivity and automatic cue approach predict unique variance in hazardous drinking. *Psychopharmacology*, *219*(2), 501-510.

Coles, M. G., Gratton, G., Bashore, T. R., Eriksen, C. W., & Donchin, E. (1985). A psychophysiological investigation of the continuous flow model of human information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(5), 529.

Critchfield, T. S., & Kollins, S. H. (2001). Temporal discounting: Basic research and the analysis of socially important behavior. *Journal of applied behavior analysis*, *34*(1), 101-122.

Dalley, J. W., Everitt, B. J., & Robbins, T. W. (2011). Impulsivity, compulsivity, and top-down cognitive control. *Neuron*, *69*(4), 680-694.

Daruna, J. H., & Barnes, P. A. (1993). A neurodevelopmental view of impulsivity. En *In: McCown WG, Johnson JL, Shure MB (eds)The impulsive client: theory, research and treatment*. Washington, D.C.: American Psychological Association.

Davidson, R. J., Putnam, K. M., & Larson, C. L. (2000). Dysfunction in the neural circuitry of emotion regulation – a possible prelude to violence. *Science*, *289*(5479), 591-594.

Dawe, S., & Loxton, N. J. (2004). The role of impulsivity in the development of substance use and eating disorders. *Neuroscience & Biobehavioral Reviews*, *28*(3), 343-351.

De Wit, H. (2009). Impulsivity as a determinant and consequence of drug use: a review of underlying processes. *Addiction biology*, *14*(1), 22-31.

Dick, D. M., Smith, G., Olausson, P., Mitchell, S. H., Leeman, R. F., O'Malley, S. S., & Sher, K.

(2010). Review: Understanding the construct of impulsivity and its relationship to alcohol use disorders. *Addiction biology*, *15*(2), 217-226.

Dickinson, A., & Dawson, G. R. (1987). Pavlovian processes in the motivational control of instrumental performance. *The Quarterly Journal of Experimental Psychology*, *39*(3), 201-213.

Donders, F. C. (1969). On the speed of mental processes. *Acta psychologica*, *30*, 412.

Duckworth, K. L., Bargh, J. A., Garcia, M., & Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological Science*, *13*(6), 513-519.

Eppinger, B., Nystrom, L. E., & Cohen, J. D. (2012). Reduced Sensitivity to Immediate Reward during Decision-Making in Older than Younger Adults. *PloS one*, *7*(5), e36953.

Eriksen, C. W. (1995). The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition*, *2*(2-3), 101-118.

Evenden, J. L. (1999). Varieties of impulsivity. *Psychopharmacology*, *146*(4), 348-361.

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, *45*(1), 152-170.

Forster, S. E., Carter, C. S., Cohen, J. D., & Cho, R. Y. (2011). Parametric manipulation of the conflict signal and control-state adaptation. *Journal of cognitive neuroscience*, *23*(4), 923-935.

Fox, E., Russo, R., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition & Emotion*, *16*(3), 355-379.

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

Gehring, W. J., Gratton, G., Coles, M. G., & Donchin, E. (1992). Probability effects on stimulus evaluation and response processes. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 198-216.

Grant, J. E., & Potenza, M. N. (2011). *The Oxford handbook of impulse control disorders*. Oxford University Press, USA.

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of general psychology*, *2*(3), 271.

Gross, J. J., & Muñoz, R. F. (1995). Emotion regulation and mental health. *Clinical Psychology: Science and Practice*, *2*(2), 151-164.

Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. *Handbook of emotion regulation*, *3*, 24.

Harmon-Jones, E., Amodio, D. M., & Zinner, L. R. (2007). Social psychological methods in emotion elicitation. *Handbook of emotion elicitation and assessment*, 91-105.

Jodo, E., & Kayama, Y. (1992). Relation of a negative ERP component to response inhibition in a Go/No-go task. *Electroencephalography and clinical neurophysiology*, *82*(6), 477-482.

Kim, S., & Lee, D. (2011). Prefrontal cortex and impulsive decision making. *Biological psychiatry*, *69*(12), 1140-1146.

Kinnison, J., Padmala, S., Choi, J.-M., & Pessoa, L. (2012). Network analysis reveals increased integration during emotional and motivational processing. *The Journal of Neuroscience*, *32*(24), 8361-8372.

Kok, A. (1986). Effects of degradation of visual stimuli on components of the event-related potential (ERP) in go/nogo reaction tasks. *Biological psychology*, *23*(1), 21-38.

Kok, A., Ramautar, J. R., De Ruiter, M. B., Band, G. P., & Ridderinkhof, K. R. (2004). ERP components associated with successful and unsuccessful stopping in a stop□signal task. *Psychophysiology*, *41*(1), 9-20.

Landis, C. (1924). Studies of Emotional Reactions. II. General Behavior and Facial Expression. *Journal of Comparative Psychology*, *4*(5), 447.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). *International affective picture system (IAPS): Technical manual and affective ratings*. Gainesville, FL: The Center for Research in

Psychophysiology, University of Florida.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual review of neuroscience*, *23*(1), 155-184.

Lhermitte, F. (1983). 'Utilization behaviour'and its relation to lesions of the frontal lobes. *Brain*, *106*(2), 237-255.

Logan, G. D. (1985). On the ability to inhibit simple thoughts and actions: II. Stop-signal studies of repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(1-4), 675-691.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological review*, *91*(3), 295-327.

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007). Time discounting for primary rewards. *The Journal of Neuroscience*, *27*(21), 5796-5804.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*(5695), 503-507.

McKAY, K. E., & Halperin, J. M. (2001). ADHD, aggression, and antisocial behavior across the lifespan. *Annals of the New York Academy of Sciences*, *931*(1), 84-96.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*, 3-19.

Miller, E. K. (2000). 22 The Neural Basis of Top-Down Control of Visual Attention in the Prefrontal Cortex. *Control of Cognitive Processes*, 511.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, *24*(1), 167-202.

Moltó, J., Montañés, S., Poy, R., Segarra, P., Pastor, M. C., Tormo Irún, M. P., … Fernández, M. C. (1999). Un método para el estudio experimental de las emociones: el International Affective Picture System (IAPS). Adaptación española. *Revista de psicología general y aplicada: Revista*

de la Federación Española de Asociaciones de Psicología, *52*(1), 55-87.

Most, S. B., Chun, M. M., Widders, D. M., & Zald, D. H. (2005). Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin & Review*, *12*(4), 654-661.

Nieuwenhuis, S., Yeung, N., Van Den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(1), 17-26.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological bulletin*, *126*(2), 220.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507-520.

Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology General*, *130*(3), 466-478.

Olmstead, M. C. (2006). Animal models of drug addiction: where do we go from here? *The Quarterly journal of experimental psychology*, *59*(4), 625-653.

Padmala, S., & Pessoa, L. (2008). Affective learning enhances visual detection and responses in primary visual cortex. *The Journal of Neuroscience*, *28*(24), 6202-6210.

Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of cognitive neuroscience*, *23*(11), 3419-3432.

Perry, J. L., & Carroll, M. E. (2008). The role of impulsive behavior in drug abuse. *Psychopharmacology*, *200*(1), 1-26.

Pfefferbaum, A., Ford, J. M., Weller, B. J., & Kopell, B. S. (1985). ERPs to response production and

inhibition. *Electroencephalography and clinical neurophysiology*, *60*(5), 423-434.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, *48*(2), 175.

Robbins, T. W., Gillan, C. M., Smith, D. G., de Wit, S., & Ersche, K. D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends in cognitive sciences*, *16*(1), 81-91.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., & Cohen, J. (2009). Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. *Journal of experimental psychology. Human perception and performance*, *35*(6), 1865.

Stautz, K., & Cooper, A. (2013). Impulsivity-related personality traits and adolescent alcohol use: A meta-analytic review. *Clinical psychology review*.

Thamotharan, S., Lange, K., Zale, E. L., Huffhines, L., & Fields, S. (2012). The role of impulsivity in pediatric obesity and weight status: A meta-analytic review. *Clinical psychology review*.

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in cognitive sciences*, *9*(12), 585.

Vuilleumier, P., & Huang, Y.-M. (2009). Emotional Attention Uncovering the Mechanisms of Affective Biases in Perception. *Current Directions in Psychological Science*, *18*(3), 148-152.

Whorwell, P. J., Houghton, L. A., Taylor, E. E., & Maxton, D. G. (1992). Physiological effects of emotion: assessment via hypnosis. *The Lancet*, *340*(8811), 69-72.

Winstanley, C. A., Theobald, D. E., Dalley, J. W., Cardinal, R. N., & Robbins, T. W. (2006). Double dissociation between serotonergic and dopaminergic modulation of medial prefrontal and orbitofrontal cortex during a test of impulsive choice. *Cerebral Cortex*, *16*(1), 106-114.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, *8*(4), 494-521.

CAPÍTULO 1. INTRODUCCIÓN

Funciones de la corteza prefrontal ventromedial en la toma de decisiones emocionales

Funciones de la corteza prefrontal ventromedial en la toma de decisiones emocionales

The role of the Ventromedial Prefrontal cortex in emotional decision making

David Contreras, Andrés Catena, Antonio Cándido, José C. Perales y Antonio Maldonado

Universidad de Granada

Correspondencia:
David Contreras
Dept. Psicología Experimental
Campus de Cartuja s/n
18071. Granada. España
e-mail: davcr@ugr.es

Published as:

Contreras, D., Catena, A., Cándido, A., Perales, J.C. & Maldonado, A. (2008). Funciones de la corteza prefrrontal ventromedial en la toma de decisiones emocionales. *Internacional Journal of Clinical and Health Psychology, 8*(1), 285-313.

Resumen

La corteza prefrontal ventromedial (VMPFC) ha sido implicada en la toma de decisiones emocionales debido a su posible participación en el aprendizaje de inversión afectivo, la propensión al riesgo y la impulsividad. Su especial entramado de conexiones con otras áreas de la corteza y con estructuras subcorticales como la amígdala justifican que pueda tener un papel de interfase entre cognición y emoción y desempeñar una función fundamental en la regulación y el control del comportamiento. En este trabajo revisamos estudios realizados con la tarea de apuestas de Iowa, tareas de aprendizaje de inversión afectivo, tareas de apuestas con diferente varianza para medir la propensión al riesgo, y tareas que introducen demora temporal de la recompensa para evaluar la impulsividad. Los datos obtenidos en esos trabajos con pacientes lesionados en la VMPFC o en otras áreas prefrontales y controles no lesionados, y datos conductuales y de actividad cerebral pueden interpretarse mejor si asumimos que la VMPFC está encargada de representar la expectativa de refuerzo. La representación de un reforzador esperado incluye la demora del reforzador, y la varianza de su magnitud.

PALABRAS CLAVE: Corteza Prefrontal Ventromedial, Emoción, Toma de Decisiones, Estudio teórico.

Abstract

The ventromedial prefrontal cortex (VMPFC) have been implied in emotional decision making, because it appear to have a role in affective reversal learning, risk proneness, and impulsive behaviour. The variety of connections of the VMPFC with several cortical and subcortical areas, as the amygdala, make likely that it can play a significant function in the control and regulation of human behaviour. In this work we reviewed studies on the Iowa Gambling task, the affective reversal learning task, lottery tasks designed to measure risk propensity, and reward delayed tasks designed to measure impulsivity. These tasks have been applied to VMPFC patients, patients with damages in other prefrontal areas, and controls. The patterns of behavioural and brain activity data can be best accounted for assuming that the VMPFC build up a representation of the expectation of reinforcement. Reinforcement amount, delay, and variance are integral parts of the representation of an expected reinforcement.

KEYWORDS: Ventromedial Prefrontal Cortex, Emotion, Decision Making, Theoretical study

Funciones de la corteza prefrontal ventromedial en la toma de decisiones emocionales

1. Introducción

El interés clínico por los mecanismos neurológicos que permiten al ser humano tomar decisiones adecuadas a sus circunstancias se remonta al menos a la segunda mitad del s. XIX, a partir del famoso caso de la lesión frontal de Phineas Gage, un empleado ferroviario al que un barreno le atravesó en 1848 la parte anterior del cráneo, penetrando por el pómulo izquierdo y saliendo por encima de la frente, con una trayectoria que afectó probablemente a la zona ventromedial pero no a la dorsolateral del lóbulo frontal (Damasio, Grabowski, Frank, Galaburda & Damasio, 1994). A pesar de que la lesión afectó de una forma tan severa el cerebro de Gage, su recuperación neurológica fue aparentemente rápida: no llegó a perder el conocimiento, y al poco tiempo desarrollaba una actividad motora normal, hablaba correctamente y entendía lo que se le decía, y no tenía problemas para recordar los acontecimientos previos y posteriores al accidente (Harlow, 1868, citado en O'Driscoll & Leach, 1998). Sin embargo, el caso se hizo célebre entre la comunidad médica de la época porque, a pesar de la ausencia de déficits habituales en otras lesiones cerebrales, en el paciente se hicieron evidentes algunas secuelas, que afectaban a facultades más abstractas de su personalidad y su comportamiento social: se volvió "caprichoso e infantil", "particularmente obstinado", con "operaciones mentales perfectas en naturaleza, pero no en grado o cantidad" (Harlow, 1868). Actualmente se considera que estos síntomas pueden estar relacionados con un déficit en la capacidad para tomar decisiones guiadas por la emoción (Clark & Manes, 2004).

La zona ventromedial de la corteza prefrontal (VMPFC) está situada en la parte ventral del prefrontal (parte de las áreas 10, 11, 12, y 47 de Broadmann, Figura 1a). Su importancia funcional es evidente si consideramos la variedad y complejidad de sus conexiones con otras áreas cerebrales (Figura 1b). En particular, es la única área prefrontal que posee densas conexiones recíprocas con la amígdala (Barbas, 2000), una estructura relacionada de forma consistente con el aprendizaje emocional, la modulación emocional de la memoria, y el reconocimiento de expresiones emocionales (Phelps & LeDoux, 2005). Además, la VMPFC es la única área prefrontal que envía un número significativo de aferencias a los centros visceromotores del hipotálamo, lo que le permite controlar la expresión autonómica de las emociones. La cara medial de la corteza prefrontal establece conexiones directas con núcleos del tronco cerebral que controlan la musculatura laríngea implicada en la fonación, lo que puede permitir a esa región cortical la modulación emocional de la expresión verbal (Barbas, 2000). La VMPFC inerva todas las demás áreas de la corteza prefrontal, a las que se atribuye un papel fundamental en funciones ejecutivas y de memoria de trabajo (Petrides, 2005). Gracias a su ubicación anatómica, la VMPFC puede encargarse de la modulación emocional de la actividad cognitiva, desempeñando un papel de intermediaria entre las estructuras cerebrales responsables de la cognición y las que controlan las emociones (Figura 1b). Por el mismo motivo, una lesión de la VMPFC puede provocar la interrupción de la principal vía de comunicación entre las áreas

emocionales y las estructuras de procesamiento cognitivo del cerebro, privando a estas últimas de información afectiva esencial para llevar a cabo su función de manera acorde a las metas del organismo.

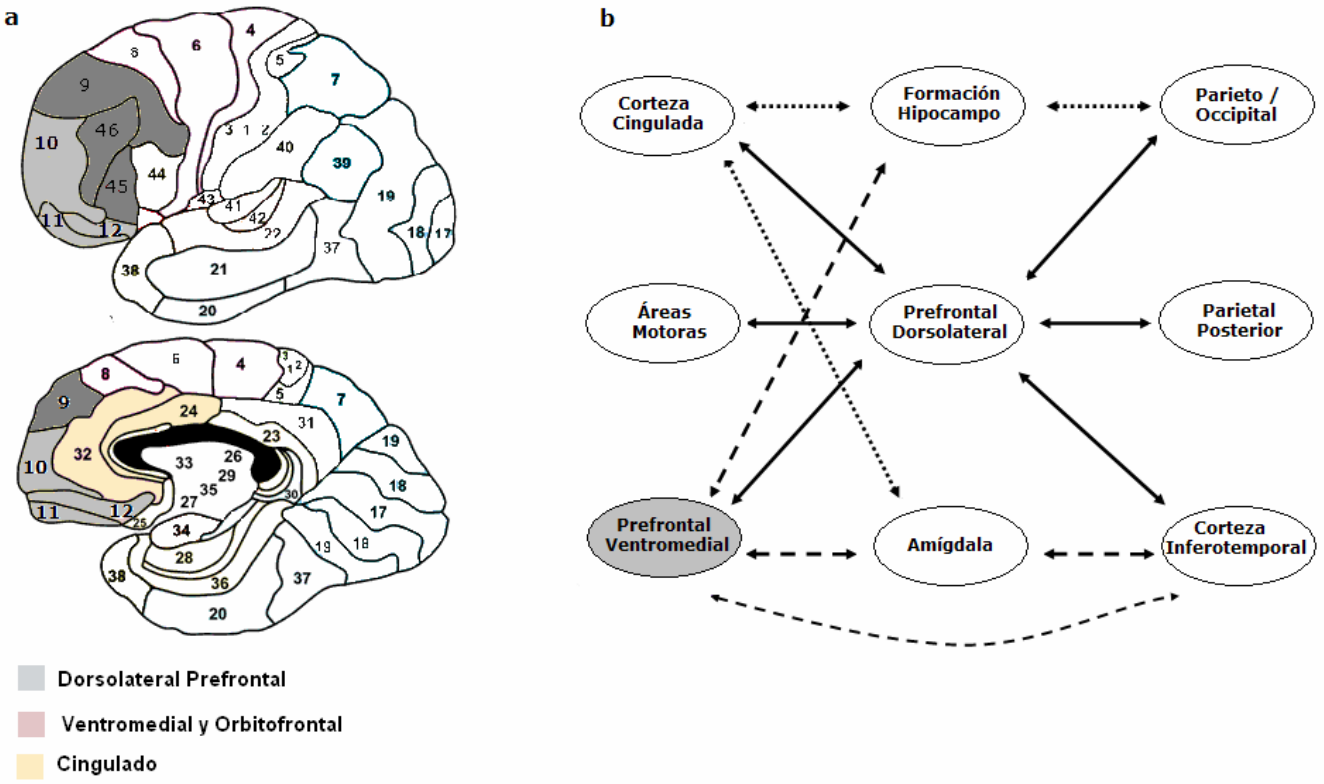

FIGURA 1. (a) Principales divisiones de la corteza prefrontal según áreas de Broadman. (b) Conexiones de la corteza prefrontal ventromedial con otras áreas corticales y subcorticales. Nótese cómo la conexión de áreas "cognitivas" (i.e.: Prefrontal Dorsolateral) con emocionales (Amígdala) se realiza a través de la corteza prefrontal ventromedial.

1.1. Síntomas de las lesiones ventromediales

Las lesiones en VMPFC producen un conjunto de síntomas que incluye: (a) dificultades en la planificación de actividades cotidianas y en la toma de decisiones; (b) alteraciones anímicas; y (c) disminución apreciable de la competencia social y deterioro de la conducta sexual. Sin embargo, permanecen relativamente intactas las capacidades ejecutivas, de resolución de problemas formales, de lenguaje, y de memoria (Dimitrov, Phipps, Zahn & Grafman, 1999).

Los pacientes lesionados suelen perder la capacidad para planificar las tareas cotidianas, administran su dinero de una forma inconsciente, que suele llevarlos a la ruina, y les es casi imposible mantener un empleo. Sus planes de futuro están pobremente especificados ycon metas poco realistas. Al mismo tiempo, son capaces de pasar horas sopesando alternativas para decidir sobre un asunto completamente trivial, como la elección de restaurante para comer (Eslinger & Damasio, 1985).

Emocionalmente, la lesión suele producir alteraciones en el temperamento: muestran más agresividad verbal, aunque normalmente no física (Grafman, Schwab, Warden & Pridgen, 1996), sienten escasa o nula empatía hacia sus semejantes, y su estado de ánimo experimenta cambios drásticos en poco tiempo. También muestran una actividad electrodermal anormalmente reducida en el visionado de imágenes de mutilaciones, pero no sexuales, algo que posiblemente esté relacionado con su reducida capacidad para empatizar (Dimitrov et al., 1999).

La falta de competencia social se manifiesta en la incapacidad de estos pacientes lesionados para mantener las amistades previas a la lesión y para formar nuevas, por la pérdida del sentido de la responsabilidad, por un comportamiento inadecuado al contexto social de cada momento (i.e.: usar lenguaje obsceno o provocativo ante personas con las que su relación es escasa), por un criterio nefasto en la elección de pareja, y por su gran vulnerabilidad al engaño y abuso por otras personas.

El reto para la investigación neuropsicológica es aclarar las dependencias funcionales entre estos déficits y caracterizar los síntomas de una forma precisa, que permita su explicación en términos de procesos cerebrales específicos. La dificultad para determinar la naturaleza exacta de los déficits provocados por las lesiones frontales se corresponde con la escasa sensibilidad de las medidas psicométricas tradicionales, tanto de personalidad como cognitivas, para detectar las alteraciones de estos pacientes, a pesar de lo evidente que resultan, para un observador cualquiera. Como consecuencia, existe en la actualidad un intenso debate sobre qué funciones fallan cuando se lesionan los lóbulos frontales, especialmente en lo concerniente a las zonas ventromedial y orbital (Fellows, 2004; Krawczyk, 2002; Kringlebach & Rolls, 2004; "The Mysterious Orbitofrontal Cortex", 2000). El interés de esta investigación es enorme, debido a que ciertos síntomas de los pacientes con lesiones ventromediales y orbitales recuerdan a los de diversas patologías que desde hace décadas se creen relacionadas con déficits neurológicos, pero en las que la localización, la naturaleza y la etiología del déficit se ignora: psicopatía (Blair, 2004; Séguin, 2004), autismo (Sabbagh, 2004), adicciones (Kalivas & Volkow, 2005; Rogers et al., 1999a), ludopatía (Brand, Kalbe, Labudda, Fujiwara, Kessler & Markowitsch, 2005), trastorno de hiperactividad y déficit atencional (Lee et al., 2005), trastorno obsesivo-compulsivo (Chamberlain, Blackwell, Fineberg, Robbins & Sahakian, 2004), o trastorno bipolar (Haldane & Frangou, 2004). Esta circunstancia ha llevado a especular con la posibilidad de que la aparición de estos trastornos esté relacionada con alteraciones neurológicas de la corteza prefrontal. La posibilidad de que el estudio neuropsicológico de esas áreas cerebrales contribuya a localizar la causa de estos trastornos y a comprender mejor su naturaleza es esperanzadora.

1.2. Organización del artículo

Este trabajo revisa los estudios que han indagado sobre los procesos cognitivos que subyacen a los déficits que padecen los lesionados ventromediales en el ámbito concreto de la toma de decisiones. Entre las hipótesis que tratan de explicar estos déficits, tres han recibido especial atención en la última década: 1) la lesión de la VMPFC produce un deterioro de la capacidad para el aprendizaje de

inversión afectivo, 2) los pacientes ventromediales son más propensos al riesgo en sus decisiones, y 3) las lesiones ventromediales incrementan la impulsividad.

En el siguiente apartado definimos brevemente los conceptos de aprendizaje de inversión afectivo, propensión al riesgo e impulsividad, y se describen algunas de las tareas que más comúnmente se han empleado para operacionalizar esos conceptos en el laboratorio. También se describe la tarea de apuestas de Iowa, muy popular en los últimos años en los estudios sobre toma de decisiones, quizá por su sensibilidad a la lesión en la VMPFC y por su validez ecológica, pero que involucra varios procesos de una forma que hace difícil analizar la contribución de cada uno. Siguen tres apartados en los que revisamos los datos a favor y en contra de la implicación de la VMPFC en el aprendizaje de inversión, en la regulación de la propensión al riesgo y en el control de la impulsividad, respectivamente. Por último discutiremos las funciones más relevantes de la corteza ventromedial en la regulación del comportamiento humano.

2. Tareas de laboratorio para la evaluación de las funciones de la VMPFC

2.1. La hipótesis del marcador somático y la tarea de apuestas de Iowa

El hecho de que los síntomas de los pacientes lesionados en la VMPFC no se manifiesten en pruebas de inteligencia tradicionales y de que estos pacientes muestren una actividad electrodermal anormal (Damasio, Tranel & Damasio, 1990) llevó a Antonio Damasio a proponer, a principios de la década de 1990, la hipótesis de que la corteza ventromedial forma parte de un mecanismo emocional cuya función es orientar al individuo en el proceso de toma de decisiones mediante la generación de estados afectivos que le informan anticipadamente de las posibles consecuencias de una acción de resultado incierto (hipótesis del marcador somático) (véase Bechara & Damasio, 2005, para revisión). La etiqueta “marcador somático” hace referencia a que las emociones anticipatorias se codifican bien en forma de ciertos cambios fisiológicos corporales, bien como una representación mental en mapas somatosensoriales de esos cambios corporales.

Con el fin de buscar apoyo empírico a su teoría, Bechara, Damasio, Damasio y Anderson (1994) diseñaron una tarea en la que intentaban reflejar las características de las situaciones de la vida cotidiana en las que tomar una decisión debería requerir de la ayuda de los marcadores somáticos: Situaciones complejas e inciertas que exigen valorar de forma intuitiva y anticipada las consecuencias futuras de las acciones, sin poseer información explícita que permita una evaluación racional. En la tarea, conocida como tarea de apuestas de Iowa (Iowa Gambling Task, IGT), el participante tiene ante sí cuatro mazos de cartas (A, B, C, y D), de los cuales debe ir extrayendo cartas, una a una, eligiendo mazo libremente. Cada carta supone una ganancia o pérdida monetaria (puntos). Dos de los mazos proporcionan ganancias altas en cada extracción (A y B), pero de vez en cuando pérdidas aún mayores, de manera que, a la larga, el resultado de extraer cartas de ellos es netamente negativo (i.e.: la pérdida neta es de 250 cada 10 extracciones). Los otros dos mazos (C y D) producen ganancias menores en cada extracción, pero las pérdidas son inferiores a las ganancias, de forma que a la larga,

extraer cartas de esos dos mazos produce un resultado neto positivo (i.e.: una ganancia neta de 250 cada 10 extracciones). Además, dos mazos tienen alta frecuencia de castigos (A y C) y los otros dos un único castigo cada 10 extracciones (B y D). El participante no tiene ningún indicio previo sobre el contenido de los mazos, por lo que debe formarse una impresión de cuáles son los buenos (tienen asociada ganancia neta) y cuáles los malos (tienen asociada pérdida neta) a lo largo del juego. Esta diseño de la tarea pretendía hacerla sensible a uno de los síntomas de los pacientes con lesión en la VMPFC, la impulsividad, la "miopía para el futuro" (Bechara, Damasio & Damasio, 2000, p.298). La idea era que la elección que proporciona mayores premios a corto plazo, aunque a largo plazo suponga una pérdida neta, debería ser más atractiva para los participantes impulsivos, en especial para los lesionados en la VMPFC. Así, los pacientes ventromediales presumiblemente sacarían más cartas de los mazos malos, "tentados" por la obtención inmediata de un premio mayor, que en los mazos buenos, sin considerar el castigo que esa decisión implicaba.

El primer estudio llevado a cabo con la IGT (Bechara et al., 1994) pareció confirmar las predicciones del grupo de Damasio. Los participantes control comenzaban sacando cartas principalmente de los mazos malos (A y B), pero, tras cierto número de ensayos, su preferencia iba cambiando hacia los mazos buenos (C y D), lo que indicaba que de alguna manera percibían que A y B resultaban netamente desfavorables a largo plazo. Por el contrario, los pacientes ventromediales no dejaban de sacar cartas principalmente de los mazos malos durante todo el juego (Figura 2).

En estudios posteriores, se observó que los controles, pero no los lesionados ventromediales, mostraban una respuesta de conductividad electrodermal (SCR) de mayor magnitud antes de sacar carta de uno de los mazos malos que antes de sacarla de uno bueno (Bechara, Tranel, Damasio, & Damasio, 1996, Bechara, Damasio, Tranel & Damasio, 1997), lo que se interpretó como una prueba de que las emociones anticipatorias, codificadas en forma de cambios corporales, ayudaban a los sujetos controles a tomar decisiones adecuadas. También se encontraron indicios de que los controles empiezan a jugar de forma ventajosa antes de ser conscientes de qué mazos son mejores y cuáles peores (Bechara et al., 1997), es decir, de que juegan correctamente de forma intuitiva. Sin embargo, Tomb, Hauser, Deldin y Caramazza (2002) observaron que los mazos malos son también los que dan premios y castigos de mayor magnitud. Utilizando una versión modificada de la IGT, demostraron que la SCR previa a la selección de un mazo está relacionada con la cuantía de los premios y castigos que proporciona el mazo escogido, no con el valor neto acumulado. Otros estudios recientes indican que el conocimiento explícito que los participantes muestran tener es suficiente para decidir ventajosamente casi desde el principio mismo de la tarea (Maia & McClelland, 2004).

Actualmente, prácticamente lo único que puede afirmarse de la IGT es que es sensible a la lesión de la corteza ventromedial (Dunn, Dalgleish, & Lawrence, 2006). No hay pruebas concluyentes de que la tarea requiera información codificada en forma de cambios corporales, ni de que las decisiones tomadas por los participantes tengan carácter intuitivo. Tampoco parece útil para medir impulsividad (véase más adelante). El rendimiento en la IGT depende de varios factores: aprendizaje

de asociación entre estímulos y reforzadores, capacidad para modificar las asociaciones previamente aprendidas cuando comienzan a aparecer castigos dentro del juego, capacidad para atender, sintetizar y recuperar de la memoria secuencias complejas de reforzamiento, y de la resolución de los conflictos de aproximación-evitación que surgen cuando un mazo está asociado simultáneamente a premio y a castigo (Fellows & Farah, 2005). Es complicado diseñar una tarea de laboratorio que sea al mismo tiempo ecológica y simple, y aunque la IGT ha sido una tarea muy popular durante los últimos diez años, está siendo reemplazada por otros paradigmas más simples (Fellows & Farah, 2003; Rolls, Hornak, Wade & McGrath, 1994) o que permiten disociar mejor diferentes procesos que pueden influir en la conducta de los participantes (Rogers et al., 1999a).

2.2. Tareas específicas para evaluar el aprendizaje de inversión afectivo

Aprendizaje de inversión (reversal learning, RL) es un término que se emplea para referirse a la capacidad de los organismos para modificar asociaciones aprendidas entre estímulos cuando cambian las relaciones de contingencia entre esos estímulos. El aprendizaje de inversión se denomina afectivo si las asociaciones que se modifican vinculan estímulos con premios o castigos. En el RL afectivo un organismo aprende que un estímulo, que anticipaba la entrega de un premio, ahora anticipa la administración de un castigo (o ya no anticipa nada), o bien que un estímulo que estaba asociado con un castigo ahora predice ahora la entrega de un premio (o ya no predice nada).

La tarea consiste en emitir o no, libremente, una respuesta de aproximación ante ciertos estímulos que se le van presentando (presionar una tecla, por ejemplo). Cada decisión que toma (presionar/no presionar) proporciona una recompensa o un castigo (i.e.: en forma de puntos o dinero agregados o retirados de su cuenta). El participante debe descubrir que es conveniente responder a unos estímulos, pero no a otros, y actuar en consecuencia. A lo largo de la tarea, los estímulos que proporcionaban recompensa pueden dejar de hacerlo o empezar a administrar castigos, y viceversa para los estímulos que proporcionaban castigos. El participante debe ajustar su conducta a la nueva situación. La medida del rendimiento del participante es el número de errores que comete cuando las contigencias cambian. Algunas variantes de esta tarea presentan los estímulos de uno en uno durante un intervalo de tiempo. El participante debe decidir responder o no al estímulo mientras está presente (tareas go-no go). Otras presentan los estímulos de dos en dos, y el participante debe seleccionar uno (discriminación, Rolls et al, 1994). En ocasiones la respuesta es verbal (Bowman & Turnbull, 2004).

2.3. Tareas específicas para la evaluación de la propensión al riesgo

Para que una decisión conlleve un riesgo, la opción escogida debe proporcionar un resultado que dependa en cierta medida del azar. Si el resultado de una de las opciones es seguro, no hay riesgo en esa opción. Una elección será tanto más arriesgada cuanta más variabilidad potencial haya en el resultado de esa elección. Esta forma de entender el riesgo proviene de la tradición de investigación sobre toma de decisiones en psicología económica (Kahneman, 2003), en la que el riesgo se define

matemáticamente como la varianza del resultado. Por esta razón, para medir la propensión al riesgo de una persona se emplean tareas en las que se le ofrecen alternativas con diferente varianza (y de igual valor esperado, normalmente, para disociar así racionalidad de propensión al riesgo), y se observa si tiende a preferir las opciones de mayor o menor varianza. Los posibles resultados de cada alternativa y las probabilidades de obtenerlos pueden comunicarse al participante explícitamente, ya sea verbal (Leland & Grafman, 2005) o gráficamente (Ernst et al., 2004; Rogers et al., 1999a), o bien se puede requerir del participante que los descubra experimentando, es decir, que los aprenda (Sanfey, Hastie, Colvin, & Grafman, 2003). En este último caso, lógicamente, el participante debe conocer el resultado final de cada decisión que toma, es decir, debe recibir feedback de ejecución, mientras que en el primer caso, el feedback no es necesario (aunque en ocasiones también se le proporciona).

2.4. Tareas para la evaluación de la impulsividad

El término impulsividad se ha empleado en la literatura científica con más de un significado (véase Evenden, 1999, para una revisión). En la tradición de investigación de psicología del aprendizaje la impulsividad se ha definido como la tendencia a preferir una recompensa inmediata y menor antes que una mayor, pero demorada. Las tareas empleadas para evaluar la impulsividad son dilemas en los que hay que escoger entre una opción que conlleva una recompensa pequeña de entrega inmediata y otra mayor, pero de entrega demorada. Las alternativas son excluyentes, es decir, escoger una implica renunciar a la otra. La información sobre las magnitudes y las demoras de las recompensas asociadas a cada alternativa puede presentarse explícitamente (McClure, Laibson, Loewenstein, & Cohen, 2004) o bien debe obtenerse experimentando (Tanaka, Doya, Okada, Ueda, Okamoto, & Yamawaki 2004). La medida de impulsividad es la tendencia del participante a preferir opciones recompensadas de forma inmediata antes que otras que reciben premios mayores pero con demora.

3. Funciones del Prefrontal Ventromedial en el aprendizaje de inversión afectivo

La capacidad para percibir un cambio en las contingencias estímulo-recompensa y estímulo-castigo y para adecuar la conducta a las nuevas circunstancias es crucial para adaptarse de forma flexible a un entorno cambiante. Un déficit en esa capacidad deja al organismo a merced de las primeras impresiones que se forme sobre cada elemento de su entorno. Una persona incapaz de llevar a cabo un aprendizaje de inversión afectivo, una vez aprendido que algo es recomendable, mantendría esa apreciación de forma indefinida, por mucha evidencia en contra que acumulara posteriormente.

Diversos estudios han implicado a la corteza prefrontal ventromedial en la capacidad para alterar asociaciones aprendidas entre estímulos y recompensas o castigos. Por ejemplo, Rolls et al (1994), empleando una tarea de discriminación visual, observaron que las personas con lesión en la VMPFC tenían dificultades para adecuar su conducta a un cambio en las relaciones de contingencia entre estímulos visuales y reforzadores. Los pacientes con lesión ventromedial cometían muchos más errores que los controles cuando las contingencias cambiaban. Igualmente cometieron numerosos

errores en otra prueba en la que a la fase de aprendizaje seguía una fase de extinción, en la que el premio se obtenía por no tocar ninguno de los dos estímulos. La incapacidad para adaptar su conducta al cambio de contingencias ocurría incluso aunque parecían conscientes de ese cambio. Por ejemplo, una participante lesionada afirmó que no iba a volver tocar la pantalla más, para hacerlo de nuevo unos pocos ensayos después. Se encontró, además, una correlación significativa (r=0.76) entre el número de errores y los síntomas de inadaptación social (evaluados personas cercanas a los pacientes). Estos resultados sugieren que la incapacidad para percibir los cambios de valor de recompensa (afectivo) de los estímulos del entorno está en la base de los síntomas de desadaptación social de las personas con lesión ventromedial (también podría haber una causa común a ambos déficits).

Por otro lado, Rolls (2000) ha propuesto que los resultados de la IGT podrían explicarse también como la consecuencia de un déficit en el RL. Dado que el programa de premios y castigos en la IGT está prefijado de tal manera que la elección de los mazos malos comienza siendo premiada con las mayores cantidades de dinero y su carácter desfavorable sólo se hace visible después de sacar muchas cartas (Bechara, Tranel & Damasio, 2000b), es probable que los pacientes ventromediales sean incapaces de modificar la conducta inicialmente aprendida de extraer cartas de esos mazos (que verdaderamente son buenos al principio). Bechara et al. (2000b) observaron, sin embargo, que las lesiones de los participantes del experimento de Rolls et al. (1994) eran laterales a las de los participantes en sus experimentos, por lo que los resultados podrían no ser comparables. Sin embargo, Fellows y Farah (2003), en un estudio en el que los sujetos del grupo de lesionados ventromediales se escogieron de una forma más cuidadosa, replicaron los resultados de Rolls, específicamente en lo concerniente a la importancia de la zona orbitofrontal, cerca del plano sagital. En otro estudio, Bowman y Turnbull (2004) diseñaron una tarea de apuestas (Bangor Gambling Task, BGT), con un solo mazo de cartas, en la que el participante debe decidir antes de cada extracción si acepta el premio o castigo de la carta siguiente o renuncia a él (go-no go con respuesta verbal). El diseño de la tarea hace que aceptar el resultado de la siguiente carta es inicialmente una elección conveniente, pero a medida que transcurre el juego, la elección más ventajosa es rechazar lo que la carta extraída pueda deparar. Se produce, por lo tanto, un cambio en la contingencia acción – recompensa: La conducta que inicialmente se premia pasa a ser castigada y viceversa. Aunque en el estudio de Bowman y Turnbull (2004) participaron solamente personas no lesionadas, sus resultados son llamativos porque la correlación entre el rendimiento en la tarea IGT y la BGT era muy alta ($r2 = 0.93$), hasta el punto que puede decirse que la IGT y la BGT, al menos para sujetos sanos, son tests paralelos. Esto vuelve a sugerir que la IGT mide fundamentalmente la capacidad para realizar un aprendizaje de inversión.

Aún más concluyente es el resultado de otro estudio de Fellows y Farah (2005), en el que se empleó una variante de la IGT en la que los primeros castigos se colocaron en posiciones iniciales de los mazos en vez de aparecer hacia la décima extracción, como ocurría en la tarea original. Esta modificación hizo desaparecer las diferencias de ejecución entre sujetos controles y lesionados ventromediales, lo que indicaba nuevamente que es el cambio en las contingencias lo que hace a la

IGT difícil para los pacientes lesionados. Si el carácter desfavorable de los mazos malos se pone de manifiesto desde el principio, los lesionados en la VMPFC no juegan peor que los controles. Además, la mejora obtenida en el rendimiento de los sujetos lesionados al cambiarse el orden de las cartas en la IGT correlacionaba apreciablemente con el número de errores que cometían en una tarea simple de RL ($r = 0.53$). Es decir, los sujetos con mayor déficit en RL eran los que más se beneficiaban del cambio en el orden de las cartas en la IGT. En el estudio de Fellows y Farah (2005) se observó, además, que los pacientes lesionados en el área dorsolateral de la corteza prefrontal (DLPFC) también eran peores que los controles no lesionados en la IGT, lo que cuestiona la especificidad anatómica de esta tarea (y limita su posible utilidad diagnóstica). Algunos estudios anteriores ya habían notado este hecho, aunque sólo cuando la lesión se producía en el hemisferio derecho (Manes et al., 2002; Clark et al., 2003). En el estudio de Fellows y Farah (2005), por el contrario, no se observaron diferencias de rendimiento entre los lesionados dorsolaterales en el hemisferio derecho y en el izquierdo. Es interesante observar que, aunque los lesionados en la la corteza prefrontal dorsolateral (DLPFC) también eran peores que los controles en la IGT, su déficit no parecía tener relación con el RL, ya que el cambio en el orden de las cartas, que bastaba para igualar a lesionados ventromediales y controles, no mejoraba el rendimiento de los lesionados dorsolaterales. Parece, por tanto, que la lesión en la DLPFC y la lesión en la VMPFC producen ambas un déficit en la IGT, pero por razones diferentes.

Respecto de la especificidad anatómica de la IGT, después de algunos estudios iniciales que implicaban de forma consistente a la corteza ventromedial pero no a otras zonas de la corteza prefrontal, otros más recientes han puesto en duda esa especificidad. Manes et al. (2002) no observaron en su estudio un deterioro en el rendimiento en la IGT en las personas con lesión en la VMPFC (aunque sí en los lesionados en la DLPFC). La razón podría ser que en el estudio de Manes et al (2002) los sujetos con lesión cerebral se clasificaron en 4 grupos, 3 de ellos con lesiones circunscritas a áreas corticales concretas (DLPFC, VMPFC o corteza prefrontal dorsomedial), y otro de personas con lesiones que afectaban a más de una de esas áreas, de manera que los sujetos del grupo de lesionados ventromediales tenía lesiones de extensión bastante reducida, mientras que en otros estudios la extensión de las lesiones no era tan limitada (Fellows y Farah, 2005). Pero también pudo influir el hecho de que, en los estudios originales de Bechara et al los participantes se seleccionaron según el criterio de que padecieran lesión ventromedial y además presentaran deterioro en la capacidad para tomar decisiones en la vida cotidiana, lo cual probablemente sesgó la muestra.

Bechara et al han argumentado en varias ocasiones en contra de que el bajo rendimiento de los pacientes ventromediales en la IGT se deba a una dificultad en el aprendizaje de inversión. En su estudio de 2000, mencionan el hecho de que, a diferencia de lo que ocurre en tareas simples de RL, como la de Rolls et al. (1994), en la que los participantes perseveran en la acción castigada, sus pacientes evitan los mazos malos tras recibir en ellos un castigo, sólo que vuelven a ellos antes y con más frecuencia que los controles. La IGT posee un componente de exploración que las tareas simples de RL no poseen. También argumentan Bechara, Damasio, Tranel y Damasio (2005) que algunos

pacientes ventromediales realizan correctamente la tarea de ordenación de cartas de Wisconsin (Wisconsin Card Sorting Task, WCST), que también mide la capacidad para detectar un cambio en las contingencias. A este respecto son especialmente esclarecedores dos estudios de Dias, Robbins y Roberts (1996, 1997) con primates no humanos, en los que se encontró una doble disociación entre el aprendizaje de inversión requerido por la WCST (detectar que ha habido un cambio en la dimensión relevante que debe ser atendida) y el requerido por una tarea como la empleada en el estudio de Rolls et al (1994), en la que lo que se modifican son las relaciones de contingencia entre varios estímulos y su valor de reforzador dentro de la misma dimensión perceptiva. En los estudios de Dias et al (1996) se demostró que el primer tipo de aprendizaje requiere de la integridad de la DLPFC pero no de la VMPFC, mientras que el segundo (afectivo) requiere la integridad de la VMPFC pero no de la DLPFC. McAlonan y Brown (2003) encontraron idéntico resultado en ratas.

El estudio original de Rolls et al (1994), y los que le han seguido sobre la función de la corteza ventromedial humana en el aprendizaje de inversión, se enmarcan en una tradición de investigaciones psicobiológicas sobre la corteza ventromedial de los primates (Rolls, 2000, para revisión). Durante tres décadas, esos estudios han permitido determinar la implicación de la VMPFC en el aprendizaje y reaprendizaje del valor afectivo de estímulos visuales, olfativos y gustativos, y en la representación del valor de reforzamiento de los estímulos según el estado motivacional del organismo (por ejemplo, la deseabilidad de una comida dependiendo de que el animal esté hambriento o no). Los estudios con humanos parecen corroborar las conclusiones obtenidas en los estudios con primates. Sin embargo, queda aún por demostrar que los problemas de adaptación social de los pacientes ventromediales puedan explicarse en términos de un simple déficit en el aprendizaje de inversión afectivo.

4. Papel de la VMPFC en la regulación de la propensión al riesgo

Otra interpretación de los síntomas producidos por la lesión de la VMPFC que se repite con frecuencia en la literatura es la de que estas personas muestran una acentuada preferencia por, o falta de precaución ante, el riesgo en sus decisiones. Desde el punto de vista estricto de la racionalidad de las decisiones, la propensión al riesgo, entendida como la preferencia por la alternativa de mayor varianza de entre dos con igual valor esperado, no puede decirse que sea una ventaja ni un inconveniente. La estrategia de optar sistemáticamente por la opción de mayor varianza producirá una secuencia de resultados con mayores fluctuaciones que la de optar siempre por la de menor varianza, pero las fluctuaciones tenderán a compensarse con el tiempo, debido a la igualdad de valores esperados. A la larga, la ganancia (o pérdida) será similar siguiendo ambas estrategias. Sin embargo, este razonamiento asume implícitamente que todas las fluctuaciones son reversibles, algo que en la realidad muchas veces no se da. Un ejemplo paradigmático es el del jugador de cartas que tras una serie de resultados negativos se queda sin blanca y debe abandonar el juego. Más en general, cualquier secuencia de resultados que conduzca a la pérdida de la vida es irreversible, y aunque nominalmente dos alternativas puedan tener igual valor esperado, si una de ellas pone en riesgo la vida, en la práctica

el valor esperado de esta última es menor. La preferencia por las alternativas arriesgadas, incluso cuando implican jugarse la vida o exponerse a la posibilidad de sufrir lesiones irreversibles, contribuye a numerosos problemas de salud pública, como los accidentes de tráfico o el tabaquismo. Por otra parte, un análisis idealizado de la propensión al riesgo puede pasar por alto el hecho de que la exposición al riesgo tiene un valor inherente porque es imprescindible para realizar conductas exploratorias que permitan mejorar las circunstancias en las que se desarrolla la vida del individuo. La exploración implica riesgo, pero sin exploración no hay posibilidad de mejora. Probablemente ésa es la razón por la que es mayor la propensión al riesgo en la conducta del adolescente que en la del adulto. Todo esto hace que la varianza de las alternativas en un dilema pueda contener valor en sí mismo, positivo o negativo, desde la perspectiva subjetiva de quien decide, hasta el punto de que una persona puede estar dispuesta a canjear varianza por valor esperado en sus estrategias de decisión, saliendo así del marco estricto de la teoría de la decisión racional.

4.1. Experimentos sin información explícita sobre las opciones

Los resultados obtenidos en las investigaciones con pacientes ventromediales empleando la IGT se han interpretado en ocasiones también como la consecuencia de un déficit en la percepción del riesgo o un sesgo hacia las decisiones arriesgadas. El motivo es que, en la IGT, los mazos malos son también los mazos de mayor magnitud de premios y castigos, por lo que es difícil discernir si la mayor preferencia por esos mazos no se debe simplemente a un sesgo hacia las elecciones arriesgadas. Así, por ejemplo, Loewenstein, Weber, Hsee y Welch (2001), en su revisión de la literatura sobre la influencia de las emociones en la toma de decisiones, hablando de la IGT afirman que "podría diseñarse fácilmente un experimento donde el valor esperado de los mazos de alto riesgo [...] sea realmente mayor que el de los mazos de bajo riesgo. En este caso, los pacientes con daño prefrontal lo harían a la larga mejor que los normales" (Loewenstein et al., 2001, p.273). En realidad, los resultados del experimento de Fellows y Farah (2005) sugieren que la preferencia por el riesgo no es el factor crucial, ya que un mero cambio de orden de las cartas en los mazos (lo cual no altera el riesgo de ninguno de ellos) elimina la diferencia entre lesionados y normales en la IGT. Sin embargo, no puede descartarse que la propensión al riesgo sea uno de los efectos de las lesiones en la VMPFC, aunque en la IGT sea difícil disociarlo de otros factores.

Sanfey et al (2003) presentaron alternativas con igual valor esperado pero diferente nivel de riesgo (diferente varianza). La tarea fue similar a la IGT, con cinco mazos de cartas con las ganancias y pérdidas distribuidas de tal manera que todos los mazos producían el mismo beneficio promedio a la larga (positivo), pero con mucha mayor variabilidad en unos que en otros. En este estudio participaron pacientes lesionados en la corteza ventromedial que fueron comparados con personas lesionadas en otras áreas de la corteza prefrontal y personas sin lesión. No encontraron diferencias entre los tres grupos en su tendencia a tomar cartas de los mazos con mayor variabilidad (todos preferían los mazos más seguros a los más arriesgados, como ocurre en general cuando todas las alternativas producen

beneficio, Kahneman, 2003). Sin embargo, mediante un análisis de cluster se observó que dentro del grupo de lesionados ventromediales podían distinguirse dos subgrupos, y uno de ellos sí mostraba una propensión al riesgo significativamente superior a los controles. De hecho, los sujetos de ese grupo preferían los mazos más arriesgados a los más seguros. Sin embargo, el estudio se llevó a cabo solamente con 8 personas lesionadas en la VMPFC, de manera que los subgrupos fueron de solo 4 miembros. Además, la secuencia de cartas dentro de cada mazo no era igual para todos los participantes, sino que se generaba aleatoriamente para cada uno de ellos. Dada la sensibilidad de los lesionados ventromediales al orden de las cartas (Fellows & Farah, 2005) es posible que las preferencias de los sujetos de ambos subgrupos difirieran solamente debido a que encontraron las cartas de los diferentes mazos en diferente orden. Además, no se halló ninguna relación entre la preferencia al riesgo de los sujetos y la localización de su lesión.

4.2. Experimentos con información explícita sobre las alternativas

Algunos equipos de investigadores han intentado superar las limitaciones de la IGT como medida de la propensión al riesgo diseñando tareas que intentan aislar la influencia de la preferencia por el riesgo de otros procesos como el aprendizaje durante la propia tarea. Rogers et al (1999a) diseñaron una tarea de apuestas (Cambridge Gamble Task, CGT) en la que al participante se le mostraba de forma explícita la información necesaria para que calculara las probabilidades de ocurrencia de los diferentes resultados posibles de cada jugada, sin tener que estimarlas a partir de su experiencia previa. De esa manera debía quedar eliminado el efecto del aprendizaje (y en particular, el del aprendizaje de inversión). El participante era situado frente a un monitor en el que aparecían 10 cajas, de colores rojo y azul, en número desigual, con un objeto amarillo escondido en una de ellas. En cada jugada debía adivinar de qué color era la caja en la que estaba escondido el objeto, y una vez escogido el color, apostar una cantidad de puntos a que su elección era correcta. Los importes posibles para la apuesta aparecían en una serie, ascendente o descendente, permaneciendo cada cantidad en pantalla 5 segundos para que el participante decidiera si escogía ésa o esperaba para escoger otra posterior. La probabilidad de que el objeto amarillo estuviera en una caja azul o una roja dependía en cada jugada solamente de la información presente en pantalla (del número de cajas de cada color), de manera que el aprendizaje no debía influir en la ejecución en la tarea. Ofrecer las apuestas en series ascendentes y descendentes permite medir de forma independiente la impulsividad motora del jugador: Un jugador impulsivo será más impaciente y tenderá a escoger una cantidad más temprana tanto en las series ascendentes como en las descendentes, de manera que para la misma jugada apostará un importe menor cuando la serie es ascendente que cuando es descendente. A mayor impulsividad, mayor diferencia en el importe para los dos sentidos de la serie. Hay que resaltar, sin embargo, que este concepto de impulsividad se refiere a la incapacidad para controlar el impulso de emitir una respuesta motora, mientras que el mencionado arriba (Bechara et al, 2000) hace referencia a la tendencia a preferir un premio menor e inmediato a uno mayor demorado. La tarea permite también medir la

propensión al riesgo del participante, medido como la cantidad que está dispuesto a apostar; la calidad de sus decisiones, observando el porcentaje de veces que escoge el color más probable (el más numeroso en pantalla); y la capacidad para ajustar el importe de la apuesta a la probabilidad de éxito, midiendo la tasa de cambio de ese importe con la proporción de cajas de un color y otro sobre el total.

Rogers et al (1999a) evaluaron a pacientes lesionados en la VMPFC y la DLPFC usando la CGT y observaron que los pacientes dañados en la VMPFC eran significativamente más lentos decidiendo su apuesta que los controles normales y lesionados. Sin embargo, exhibían una menor preferencia por el riesgo, ya que sus apuestas, ante la misma jugada, eran menores que las de los lesionados en la DLPFC y los controles no lesionados. No se observó impulsividad motora en ninguno de los grupos. Manes et al (2002) examinaron también a los participantes en su estudio con la CGT, y, en lo referente a la propensión al riesgo, sólo los participantes con lesiones que abarcaban más de una zona en la corteza prefrontal apostaron cantidades significativamente mayores que los controles no lesionados. Los otros tres grupos de lesionados (sólo lesión dorsomedial, sólo dorsolateral y sólo ventromedial) no mostraron diferencias significativas entre sí ni con los controles. Ninguno de los grupos fue significativamente más lento que los demás en la deliberación sobre su apuesta. Otros dos estudios con la CGT en los que los participantes padecían la variante frontal de la demencia frontotemporal (Rahman, Sahakian, Hodges, Rogers & Robbins, 1999) y hemorragia subaracnoidea de la arteria comunicante anterior (Mavaddat, Kirkpatrick, Rogers & Sahakian., 2000) sí pusieron de manifiesto una mayor propensión al riesgo en estos pacientes, sin embargo, no se comprobó la especificidad anatómica del déficit neurológico. No parece, por tanto, que los estudios que han empleado la CGT hayan demostrado de forma clara una relación directa entre la lesión en la VMPFC y la propensión a asumir riesgos, a pesar de que las observaciones clínicas sugieren esta relación. Aunque, obviamente, un resultado nulo no equivale a demostrar que la relación no existe.

Rogers et al (1999b) adaptaron la CGT para su administración durante pruebas de neuroimagen. La versión modificada sólo presentaba en la pantalla 6 cajas de color rojo y azul, en proporciones 5:1, 4:2 o 3:3. En lugar de ofrecérsele al participante cantidades a apostar en una serie, se le mostraban en pantalla solamente dos alternativas, una para la apuesta por un color y otra para la apuesta por el otro, siempre una cantidad mayor para el color menos numeroso. Por ejemplo, con cinco cajas rojas y una azul, se le podía dar al jugador opción de apostar bien 30 puntos al color rojo o bien 70 al color azul. Llama la atención el hecho de que ofrecer la posibilidad de apostar un importe mayor a la alternativa menos probable (el azul en este ejemplo) hacia más atractiva esa opción, a pesar de que la apuesta por el color menos frecuente tiene un valor esperado negativo (es más probable perder que ganar), y el valor esperado es tanto más negativo cuanto mayor sea la cantidad en juego. Por ejemplo, si con cuatro cajas rojas y dos azules las apuestas posibles eran 10 al rojo y 90 al azul, era mayor el número de participantes que elegían apostar al azul que cuando las apuestas posibles eran 30 y 70, algo totalmente en contra de lo predicho por la teoría de la decisión racional. Lamentablemente, en el diseño de la prueba están confundidos la propensión al riesgo con la

racionalidad de la decisión tomada (mayor riesgo implica también valor esperado negativo), de forma que en el análisis de los resultados es imposible disociar ambos factores. Rogers et al (1999b) observaron que la actividad cerebral (flujo sanguíneo medido mediante fMRI) era mayor mientras los participantes tomaban decisiones en la tarea de apuestas que en la condición de control. Sin embargo, no realizaron la sustracción entre los casos en los que el sujeto escogió la opción más probable y los casos en los que escogió la menos probable, de manera que sólo probaron la implicación de la VMPFC en la resolución del conflicto entre riesgo y recompensa, pero no específicamente en la selección de la opción menos (o más) arriesgada.

Un análisis de este tipo fue realizado por Ernst et al (2004) empleando una tarea llamada "rueda de la fortuna" (Wheel Of Fortune, WOF). En esta prueba se presenta al sujeto un círculo con un sector circular coloreado de azul y el resto de color magenta. Sobre el círculo debe imaginar que gira una ruleta que termina deteniéndose sobre la parte magenta o azul. Se le ofrece la posibilidad de apostar una cantidad al color azul o bien otra cantidad diferente al magenta. La proporción de áreas entre los dos colores oscila entre 1:7 y 4:4 y las cantidades ofrecidas para apostar son inversamente proporcionales al área ocupada por el color correspondiente, de manera que si un color ocupa un área pequeña el participante puede apostar por él una cantidad grande, y si ocupa un área grande, puede apostar una cantidad pequeña. A diferencia de lo que ocurre en la CGT, si el jugador no acierta, no pierde la cantidad apostada, sino que simplemente no la gana, es decir, si acierta, gana lo apostado, pero si falla no pierde nada. Así, apostar por la alternativa menos probable, pese a ser menos probable, tiene un valor esperado positivo, y dado que probabilidades y magnitud de la apuesta son inversamente proporcionales, el valor esperado de ambas alternativas es prácticamente el mismo. De ese modo, escoger una u otra alternativa es estrictamente una cuestión de preferencia por el riesgo, ya que apenas se mezcla riesgo y racionalidad de la decisión. Ernst et al (2004) compararon la actividad cerebral de los sujetos cuando optaban por la alternativa más segura con la actividad cuando escogían la más arriesgada, y observaron que, en el caso de escoger la opción arriesgada, la corteza prefrontal ventral (orbitofrontal) mostraba mayor actividad bilateralmente, aunque de forma más acentuada en el hemisferio derecho. Por tanto, al parecer, la VMPFC interviene en los sujetos no lesionados cuando optan por escoger una alternativa con mayor riesgo frente a una con un valor esperado comparable pero menos arriesgada. Este hecho podría estar relacionado con la mayor propensión al riesgo que los informes clínicos atribuyen a las personas con lesión ventromedial.

4.3. Experimentos sin feedback de ejecución

En la literatura económica, una situación azarosa en la que se conocen con exactitud las probabilidades de los diferentes posibles desenlaces se denomina situación "de riesgo". "Incertidumbre", en cambio, es un término que se reserva para las situaciones en las que, además de intervenir el azar, se desconocen las probabilidades de los posibles resultados. La IGT, o la tarea del estudio de Sanfey et al (2003) son tareas de incertidumbre más que de riesgo, al menos en las primeras fases de la prueba,

dado que el sujeto ignora por completo con qué se va a encontrar en los distintos mazos. Eliminar el efecto del aprendizaje equivale, usando esta terminología, a convertir una tarea de incertidumbre en una tarea de riesgo. De ahí la importancia de hacer explícita (o al menos, accesible por medio de un cálculo sencillo) la información sobre las probabilidades de los diferentes resultados de cada escenario de decisión, si lo que se desea es medir la propensión al riesgo del participante. Sin embargo, hacer accesible esa información no garantiza en modo alguno que el sujeto haga uso de ella. En los experimentos de Rogers et al (1999a, 1999b), se dio por sentado que explicitar la información suficiente para calcular las probabilidades bastaba para que el sujeto no se apoyara en el aprendizaje al tomar decisiones, pero los propios resultados del experimento hacen sospechar que no era así exactamente (por ejemplo, en el primer estudio realizado con la CGT, Rogers et al, 1999a, los sujetos con lesión ventromedial mejoraron significativamente la calidad de sus decisiones entre dos ejecuciones sucesivas de la misma tarea, escogieron la opción más probable en un 81% de las ocasiones la primera vez que realizaron la tarea, frente a un 89% la segunda vez). En la versión de la CGT adaptada para pruebas de neuroimagen (Rogers et al, 1999b), el propio diseño del experimento hacía engañosa la información visible en pantalla, ya que se acopló la condición en la que la proporción de cajas rojas a azules era 5:1 a la condición en la que la razón era 4:2, de modo que el número de aciertos fuera el mismo en ambas condiciones. Eso equivale a alterar artificialmente las probabilidades de que el objeto escondido estuviera en las cajas de ambos colores, algo que posiblemente los participantes pudieron detectar de forma implícita. En cualquier caso, aún sin detectar el cambio en las probabilidades, nada garantiza que los sujetos no sospechasen que la tarea estaba trucada, desconfiasen de la información presente en pantalla y tratasen de extraer conclusiones por su cuenta. El único modo de garantizar que el aprendizaje no influye en el rendimiento de los participantes es no proporcionarles información alguna sobre el resultado de sus decisiones. Es decir, eliminar el feedback de ejecución.

Leland y Grafman (2005) midieron la propensión al riesgo de pacientes con lesión en la VMPFC mediante un cuestionario similar a los que se han empleado durante décadas en los estudios de toma de decisiones en psicología económica (Kahneman, 2003). Cada ítem describe dos juegos de azar (loterías) y pregunta al participante a cuál de ellos preferiría jugar. Todos los ítems muestran de forma explícita las probabilidades de cada posible resultado, de manera que la tarea es de riesgo, no de incertidumbre. Para disociar la calidad de la decisión tomada de la propensión al riesgo, los ítems del cuestionario de Leland y Grafman (2005) ofrecían siempre dos loterías con el mismo valor esperado, pero diferente varianza. En este estudio no se encontró ninguna diferencia entre el grupo de pacientes con daño ventromedial y los controles en la preferencia por las loterías de mayor varianza. Es decir, enfrentados con dilemas como el del ejemplo anterior, los pacientes ventromediales no aparentan ser más propensos al riesgo que los sujetos normales. (De hecho, el estudio midió también la impulsividad y la adecuación de las respuestas sociales de los sujetos mediante cuestionario, y no encontró diferencia alguna entre sujetos normales y lesionados). Leland y Grafman (2005) atribuyeron este

resultado nulo a que, en su estudio no se proporcionaba feedback de ejecución a los participantes. Compararon sus resultados con los de Rogers et al (1999a), con la CGT, el original de Bechara et al (1994), con la IGT, un estudio de Breiter, Aharon, Kahneman, Dale y Shizgal (2001), con un paradigma similar al de la rueda de la fortuna, y otros dos estudios, uno de Critchley, Mathias y Dolan (2001), en el que se presentaban a los sujetos cartas al azar, numeradas del 1 al 10, y éstos debían adivinar si la siguiente carta tendría un número mayor o menor, y otro de Rustichini, Dickhaut, Ghirardato, Smith y Pardo (2005) con una metodología similar a la suya (dilemas sobre loterías). Salvo su estudio y el de Rustichini et al (2005), todos los demás sugerían un papel de la VMPFC en la toma de decisiones (bien porque la lesión ventromedial empeoraba el rendimiento, o bien porque se observaba activación de dicha zona mediante neuroimagen). Comparando los estudios que sí indicaban implicación de la VMPFC con los que no, el único factor que permitía distinguir unos estudios de otros era el feedback de ejecución. Otros factores como la necesidad de recurrir al aprendizaje, la promesa de incentivos monetarios, la presentación de la información sobre probabilidades de forma explícita o implícita, o la existencia de incertidumbre, no permitían diferenciar entre los trabajos que sugerían un papel de la VMPFC y los que no.

Para Leland y Grafman, la eliminación del feedback de ejecución privaba a la tarea de su carácter emocional (la tarea se volvía "aburrida"), lo que explicaba su independencia de la integridad de la VMPFC. De este modo, sería imposible diseñar una tarea de toma de decisiones emocionales en la que se garantizara que el aprendizaje no interviniera mediante la eliminación del feedback. Sin embargo, este análisis ignora los resultados de otro experimento (Dickhaut, McCabe, Nagode, Rustichini, Smith & Pardo, 2003), en el que sí se observó activación ventromedial a pesar de que no se proporcionó a los sujetos feedback de ejecución. Como en el experimento de Leland y Grafman (2005), en el estudio de Dickhaut et al (2003) se ofrecían a los participantes loterías igualadas en valor esperado, pero de diferente varianza, para que escogieran la que consideraran preferible. Del mismo modo que en el estudio de Leland y Grafman (2005), y a diferencia de lo que ocurre en la IGT o la CGT, los participantes no podían comprobar lo acertado de su decisión de forma inmediata, ya que la lotería no se jugaba. Sin embargo, y en esto el diseño difería del de Leland y Grafman (2005), al participante se le advertía que dos de las loterías que eligiera durante la prueba se seleccionarían (al azar) y se jugarían al final de la misma, con dinero real. Con este diseño (prácticamente idéntico, en lo demás, al empleado por Rustichini et al, 2005), se observó que un área extensa de la VMPFC mostraba actividad significativamente superior cuando el sujeto tenía que decidir entre dos loterías con valor esperado positivo que cuando tenía que optar entre dos que implicaban pérdidas, pero solamente en el caso de que ninguna de ellas tuviera varianza nula (figura 5). Es decir, la diferencia de activación entre decisiones que implicaban pérdidas y ganancias sólo se daba cuando ninguna de las alternativas carecía de riesgo. De ese modo, la presencia de riesgo en la situación de decisión hacía necesario el reclutamiento de la corteza ventromedial dependiendo del signo de las cantidades en juego.

El hecho de que a lo largo de toda la tarea no se de al participante ninguna información sobre la calidad de sus decisiones garantizaba con certeza (a diferencia de lo que ocurría en el experimento de Rogers et al, 1999a), que la actividad de la corteza ventromedial no tenía ninguna relación con el aprendizaje instrumental. Por otro lado, la razón por la que el paradigma de Leland y Grafman (2005) no permite discriminar entre lesionados y normales probablemente tenga más relación con la anticipación del reforzador que con el hecho de que el feedback se reciba de forma inmediata, a continuación de cada decisión tomada. En el experimento de Leland y Grafman (2005) los participantes no se jugaban nada, por así decirlo. La relevancia hedónica de las decisiones tomadas era nula, puesto que, ni de forma inmediata, ni al final del experimento, nada de lo que decidieran tendría ninguna consecuencia para ellos mismos (y probablemente era esto, más que el feedback, lo que hacía a la tarea "aburrida"). En cambio, en la IGT y otros paradigmas en forma de juego, aunque no se apueste con dinero real, la mera motivación de logro probablemente es suficiente aliciente para que el resultado de las acciones se experimente como un premio o castigo. Del mismo modo, en el estudio de Dickhaut et al (2003), cualquiera de las decisiones tomadas durante la prueba por los sujetos podía implicar un castigo o una recompensa al final de la misma. Podría ser, por tanto, la anticipación del castigo o recompensa lo que hacía diferente el estudio de Leland y Grafman (2005) de los demás. Esta interpretación de los resultados, además, concuerda con la abundante evidencia de que la VMPFC y el núcleo basolateral de la amígdala forman parte del circuito neuronal que se encarga en la conducta instrumental de la representación de una recompensa anticipada como consecuencia de una acción (véase Holland & Gallagher, 2004, para revisión).

En resumen, para analizar específicamente la propensión al riesgo es necesario igualar el valor esperado de las alternativas sobre las que el participante ha de decidir, ofreciéndole, claro está, alternativas de diferente varianza. Así se elimina la confusión entre propensión al riesgo y racionalidad de las decisiones. No proporcionar feedback de ejecución asegura que la tarea no depende del aprendizaje, y permite discriminar el efecto del riesgo del efecto de la incertidumbre. Sin embargo, despojar a la tarea de toda relevancia hedónica para los sujetos anula el carácter emocional de la toma de decisiones, por lo que la eliminación del feedback no debe suponer que las decisiones tomadas no tengan ninguna consecuencia para el sujeto. Una forma de conseguir esto es diferir las consecuencias de las decisiones tomadas hasta el final de la prueba. En esas condiciones la VMPFC parece desempeñar cierto papel, por lo que se observa en pruebas de neuroimagen. Sin embargo, no está claro si la lesión de esta área cerebral produce propensión al riesgo, como sugieren las observaciones clínicas, o, por el contrario, produce aversión. Tampoco está claro si su papel en la propensión al riesgo depende del signo del valor esperado de las opciones sobre las que hay que decidir.

4.4. Experimentos con animales

La función de la VMPFC también se ha investigado en animales lesionados. En este caso, puesto que la información sobre las alternativas no puede ser explícita, es prácticamente inevitable que se

mezclen efectos de aprendizaje. En un experimento con ratas, Mobini et al (2002) observaron que la lesión de la VMPFC (previa al entrenamiento) producía en los animales aversión al riesgo. Las ratas aprendían la relación entre presionar una de dos palancas (A y B) y número de bolitas de comida (1 y 2, respectivamente para A y B). La probabilidad de que la pulsación de B produjera resultado se redujo progresivamente de 1 a 0, sin alterar la de la otra palanca. Las ratas lesionadas en la VMPFC eran más sensibles que las controles al cambio de contingencias: Mostraron preferencia por la opción segura y de menor valor (A) mucho antes que las ratas no lesionadas. Por ejemplo, cuando pulsar la palanca B sólo era efectivo en un 20% de los casos, las ratas no lesionadas seguían escogiendo esa opción cerca de un 50% de los ensayos, mientras que las lesionadas sólo la escogían en un 20% de las ocasiones. Esta conducta puede interpretarse como una menor propensión al riesgo en las ratas lesionadas que en las intactas. En el diseño, sin embargo, la varianza del resultado y el valor esperado estaban confundidos. También se mezclaban efectos del aprendizaje, y aunque un fallo en el aprendizaje de inversión debería haber hecho a las ratas lesionadas menos sensibles al cambio de las contingencias que las no lesionadas, la influencia del aprendizaje complica la interpretación de los resultados.

En síntesis, los resultados de los estudios anteriores indican que la corteza ventromedial está implicada en las decisiones con relevancia hedónica que implican alternativas con riesgo. Sin embargo, parece claro que su papel no puede describirse diciendo simplemente que es un regulador de la propensión o aversión al riesgo entre alternativas de igual valor esperado.

5. Papel de la VMPFC en la regulación de la conducta impulsiva

El término impulsividad se emplea para referirse a un conjunto amplio de fenómenos, y con frecuencia se ha usado en la literatura de un modo poco preciso. Dentro de la tradición de la psicología del aprendizaje, la conducta impulsiva es la tendencia a preferir un premio menor inmediato a otro mayor pero demorado. En ese contexto, “impulsiva” es, por tanto, un adjetivo aplicable a la conducta de elección. Esta definición es fácil de operacionalizar, pero tiene el inconveniente de asumir que la impulsividad es un rasgo monolítico de la conducta, que afecta por igual a toda decisión independientemente del tipo de estímulos implicados y del escenario donde se decide, cuando quizá la tendencia a preferir un premio inmediato menor no se transfiera de unas categorías de estímulos a otras o de unas situaciones a otras. Además, el empleo del verbo “preferir” en la definición parece sugerir que la elección del premio menor e inmediato es el resultado de una comparación entre las dos alternativas, sin embargo, la observación del resultado (la alternativa elegida) no permite deducir que ése haya sido el proceso. La elección del premio inmediato y menor puede deberse simplemente a que la persona responde automáticamente ante esa opción, sin llegar, en sentido estricto, a compararla con la otra. Es interesante que en el lenguaje cotidiano el término “impulsividad” se emplee como sinónimo de “actuar sin pensar”, lo que concuerda con esta segunda interpretación de la preferencia del individuo impulsivo. Esta otra acepción de la impulsividad está más relacionada con la ejecución

de la respuesta que con la valoración del resultado. Evenden (1999) diferencia entre una impulsividad relacionada con el resultado (incapacidad para soportar una demora de la recompensa) y una impulsividad en la respuesta (dificultad para controlar una respuesta automática ante un estímulo), que corresponden básicamente a las dos acepciones mencionadas del concepto de impulsividad. También distingue otro tipo de impulsividad relacionada con la preparación de la respuesta (tendencia a responder antes de haber adquirido la información necesaria para hacerlo de forma correcta). Estos tres tipos de impulsividad se han disociado en experimentos con fármacos serotoninérgicos en ratas (Evenden, 1999), por lo que probablemente dependen de sistemas cerebrales diferentes.

La organización de la actividad humana renunciando al impulso de obtener un beneficio inmediato para asegurarse la obtención de un mayor beneficio futuro es probablemente el tema central de la economía como disciplina científica. Toda la economía gira en torno al concepto de capital, el producto de la actividad humana que no se emplea para el consumo inmediato, sino como medio para la producción futura o para el consumo en épocas posteriores de escasez. El control de la impulsividad es la característica de la conducta humana que permite a nuestra especie asegurarse la supervivencia a largo plazo atenuando los efectos de las fluctuaciones climáticas o de los desastres naturales. Por otro lado, en el ámbito de la conducta individual, impulsividad (o su control) es un concepto necesario para explicar patrones de conducta como el sacrificio sistemático del atleta que entrena durante meses para una sola prueba, o el del opositor que prepara durante años un solo examen. En psicología clínica, toda una línea de investigación sobre las causas de las adicciones relaciona la conducta adictiva con un fallo en el control de la impulsividad. Desde este punto de vista, se considera que el adicto lo es porque tiende a ceder ante el placer inmediato que le proporciona su adicción a pesar de las consecuencias nefastas que esta adicción conlleva a largo plazo (Robinson y Berritge, 2003).

El control de la impulsividad, de forma genérica, tiene relación con el concepto de autorregulación, de inhibición de tendencias de acción primarias o automáticas. Dentro del sistema nervioso, las estructuras filogéneticamente más recientes se encargan en general de inhibir las tendencias de respuesta de las más antiguas. No es extraño, pues, que la lesión de la VMPFC provoque efectos que globalmente puedan describirse como desinhibición conductual. Sin embargo, es improbable que el efecto de una lesión ventromedial en la impulsividad se pueda resumir diciendo simplemente que se ha alterado la sensibilidad a la demora de los reforzadores, de un modo indiscriminado.

5.1. Investigaciones con animales

Los experimentos con animales en los que se ha investigado el efecto de las lesiones de la VMPFC sobre la impulsividad han obtenido resultados contradictorios, al menos en apariencia. Mobini et al (2002), aumentaron progresivamente el intervalo de demora entre la pulsación de la palanca B y la entrega del reforzador correspondiente, sin variar el intervalo de demora para la palanca A, y observaron que las ratas lesionadas empezaban a encontrar preferible la opción A (recompensa menor

inmediata) frente a la B (mayor y demorada) mucho antes que los controles. Es decir, la lesión ventromedial inducía impulsividad en los animales, en concordancia con las observaciones clínicas en humanos. Sin embargo, las diferencias de procedimiento (lesión antes del aprendizaje versus después del mismo, manipulación de los intervalos intra o entre sesiones), hace creíble que en el experimento de Winstanley et al (2004), pero no en el de Mobini et al (2002), el aprendizaje de inversión afectivo intrasesión desempeñara un papel significativo, puesto que en cada sesión la rata debía aprender que la recompensa de mayor magnitud que se obtenía presionando una de las palancas, que inicialmente era inmediata, posteriormente se recibía con una demora, que iba creciendo a lo largo de la sesión. Quizá la menor sensibilidad de las ratas a la demora del reforzador, en el experimento de Winstanley et al (2004), se debiera más bien a la dificultad para actualizar el valor del reforzador con la nueva demora que a la mayor preferencia en sí por un reforzador demorado.

5.2. Experimentos con neuroimagen

La IGT se diseñó inicialmente con el propósito de ser sensible a la impulsividad del participante (a su incapacidad para soportar una demora de la recompensa). Maia y McClelland (2004), sin embargo, rechazaron que la IGT midiera la impulsividad de los participantes. Para ellos es claro que en la IGT, debido a su propio diseño, cada decisión tomada por el sujeto tiene como única consecuencia la pérdida o ganancia de la cantidad asociada a la carta que se extrae con esa decisión. Las pérdidas o ganancias futuras son consecuencia de decisiones posteriores. Carece de sentido, por lo tanto, decir que el sujeto que saca carta de un mazo malo está cediendo a la tentación de un beneficio inmediato a pesar de que eso acarrea pérdidas posteriores, ya que la decisión de escoger un mazo malo acarrea como única pérdida la que se pueda obtener de la propia extracción realizada. Esto, desde luego, no implica que los pacientes de lesiones en la VMPFC no puedan mostrar una tendencia a la conducta impulsiva, pero, sea así o no, el hecho es que la IGT no aclara nada sobre esa cuestión.

Algunos experimentos de neuroimagen han evaluado la toma de decisiones en las que el participante debe simplemente escoger entre dos premios de diferente magnitud y diferente demora. En ellos se ha observado una tendencia de las áreas ventromediales de la corteza prefrontal a activarse en las decisiones que suponen la recepción inmediata de reforzadores, mientras que en las decisiones que conllevan la demora del reforzador las áreas prefrontales más implicadas son dorsales y laterales. McClure et al (2004) registraron la actividad cerebral mientras los participantes escogían entre vales de compra a canjear de inmediato (en el mismo día) y otros, de mayor importe, canjeables solamente unas semanas después. Antes de comenzar, se avisaba a los participantes de que una de las elecciones efectuadas por ellos se escogería al azar al final del experimento y se les haría entrega del vale correspondiente. Mediante regresión a un modelo lineal de dos factores, uno relacionado con las elecciones inmediatas y otro con el proceso de elección en general, observaron que las áreas más activas en la selección de alternativas de resultado inmediato eran la VMPFC, el estriado ventral, y la corteza prefrontal medial, mientras que las áreas relacionadas de forma general con el proceso de

decisión se encontraban en el surco intraparietal de ambos hemisferios, y en el hemisferio derecho en la DLPFC, corteza prefrontal ventrolateral y corteza orbital lateral.

Un resultado similar obtuvieron Tanaka et al (2004) en otro experimento en el que los participantes aprendían a jugar un juego en el que se alternaban dos fases, una en la que las decisiones podían tomarse guiándose únicamente por la búsqueda de un beneficio inmediato, ya que no tenían repercusiones posteriores, y otra en la que, para poder obtener beneficios a largo plazo, era necesario tomar algunas decisiones desfavorables a corto plazo. La sustracción de la actividad cerebral de ambas fases era, dentro del lóbulo frontal, mayor en la DLPFC, corteza prefrontal ventrolateral e ínsula, cuando el sujeto tenía que tomar decisiones desfavorables a corto plazo como medio para lograr beneficio a largo plazo. Es decir, esas zonas, presumiblemente, están relacionadas con la representación de un reforzador demorado. También se observó gran activación en la zona dorsal del núcleo estriado. Por el contrario, la comparación entre la activación en la fase del juego en la que se podían tomar decisiones teniendo en cuenta sólo los resultados inmediatos y otra condición en la que no se obtenía ningún reforzador mostró que, dentro del lóbulo frontal, la activación era mayor en el primer caso en el área lateral de la corteza orbitofrontal, y en la ínsula. También era mayor la activación en la zona ventral del núcleo estriado. Estas activaciones se suponen relacionadas con la representación de un reforzador inmediato. Es interesante observar que la activación en la zona dorsal del núcleo estriado que se producía en relación con la representación de un reforzador demorado se acompañaba de activación en regiones corticales con numerosas conexiones recíprocas con esa área del núcleo estriado, y del mismo modo ocurría con el patrón de actividad relacionado con la representación de un reforzador inmediato, que implicaba a la zona ventral del estriado y a áreas corticales con abundantes conexiones recíprocas con esa zona. Estos resultados sugieren que el procesamiento de reforzadores inmediatos activa ciertos bucles cortico-basales mientras que el de reforzadores demorados activa otros diferentes. Aunque esta tarea mezcla la toma de decisiones en sí con el aprendizaje de contingencias estímulo-respuesta-reforzador, los resultados de este estudio convergen con los de McClure et al (2004) en el sentido de implicar al bucle que une al estriado ventral con la corteza prefrontal ventral en la representación de reforzadores inmediatos.

5.3. Experimentos con pacientes con lesión

Otros experimentos han explorado mediante medidas conductuales y de autoinforme la impulsividad de pacientes con lesiones en la VMPFC. Berlin, Rolls y Kischka (2004) obtuvieron medidas simultáneas de diversos aspectos de la impulsividad en lesionados en la corteza orbitofrontal (OFC), empleando como controles personas sin lesión y personas con una lesión prefrontal que no afectaba a la OFC. Como medidas de impulsividad emplearon un cuestionario de autoinforme (Barratt Impulsive Scale, BIS), una prueba conductual en la que el sujeto debía señalar de entre 12 figuras muy parecidas cuál era idéntica a otra de muestra (Matching Familiar Figures Test, MFFT), y cuatro tareas de percepción del transcurso del tiempo. Los cuestionarios de autoinforme como medida de impulsividad

proceden de una tradición de investigación psicométrica que trataba de desarrollar una teoría de la personalidad en la que la impulsividad pudiera identificarse como un rasgo (Evenden, 1999). El test MFFT fue creado por Kagan (1966) como medida de la "impulsividad en la reflexión", concepto que coincide con la impulsividad en la preparación de la respuesta que propone Evenden (1999). La idea de emplear una medida de la percepción subjetiva del transcurso del tiempo como modo de evaluar la impulsividad se basa en la hipótesis de Barratt (1983) de que las personas para las cuales el tiempo subjetivo transcurre más deprisa percibirán como menos frustrante la demora de un reforzador, y, en consecuencia, tenderán a ser menos impulsivos, mientras que ocurrirá lo contrario con aquellas personas para las que el tiempo subjetivo transcurre más despacio. En el estudio de Berlin et al (2004) se utilizaron cuatro tareas para analizar la percepción subjetiva del transcurso del tiempo en los participantes: Una de producción temporal en la que el sujeto debía indicar cuándo consideraba que había transcurrido un intervalo de tiempo de un cierto número de segundos; otra de estimación temporal en la que debía adivinar cuántos segundos había durado un intervalo cuyo principio y final indicaba el experimentador mediante una señal; otra de tasa temporal en la que debía contar en alto intentando matener un ritmo constante de un número por segundo; y otra de estimación a largo plazo, en la que tenía que responder a la pregunta "¿cuánto tiempo ha transcurrido desde el principio de la prueba?".

La prueba de autoinforme contenía tres subescalas, una de impulsividad motora (de respuesta), otra de impulsividad en la planificación (que correspondía aproximadamente a la impulsividad en la reflexión de Kagan, 1966) y otra de impulsividad cognitiva (sensibilidad a la demora del resultado). Los sujetos lesionados en la OFC puntuaron significativamente más alto que los controles en impulsividad motora, en impulsividad en la planificación y en la escala global, pero no en impulsividad cognitiva. En la MFFT, los lesionados OFC fueron más rápidos al responder y cometieron más errores que los controles, es decir, mostraron más impulsividad en la reflexión. Y en las pruebas de percepción temporal, los lesionados orbitofrontales producían intervalos de tiempo real más cortos que los controles cuando se les pedía un intervalo de un cierto número de segundos, mientras que cuando debían estimar el número de segundos transcurridos en un intervalo de tiempo dado, sus estimaciones eran muy superiores. Ambas medidas indican consistentemente que la lesión orbitofrontal produce una aceleración del tiempo subjetivo, lo cual podría inducir impaciencia en la espera de un reforzador demorado, y, por tanto, impulsividad. Conjuntamente, tanto las medidas de autoinforme como las conductuales de este estudio sugieren que la lesión de la OFC induce impulsividad, al menos en los aspectos de preparación y producción de la respuesta motora.

6. Conclusiones

La VMPFC desempeña un papel crucial en el proceso de toma de decisiones sobre asuntos cuyas consecuencias potenciales afectan directamente a la persona. La implicación de la VMPFC está demostrada en el proceso de actualización de preferencias (reversal learning; Rolls et al, 1994), y hay

indicios claros de su intervención durante la expresión de esas preferencias en la toma de decisiones cuando las alternativas sobre las que se decide implican riesgo (Dickhaut et al, 2003; Rogers et al, 1999b) o proporcionan reforzadores con diferente demora (McClure et al, 2004), aunque en este caso su función no está clara.

La lesión de la VMPFC reduce drásticamente la capacidad de las personas para llevar a cabo aprendizaje de inversión afectivo (Fellows & Farah, 2003; Rolls et a., 1994), y posiblemente este déficit sea el responsable de muchos de sus problemas de adaptación social (Rolls et al, 1994). Así mismo, el fallo en el RL afectivo probablemente es suficiente para explicar el bajo rendimiento en la IGT de los lesionados ventromediales, pero no el de los dorsolaterales (Fellows & Farah, 2005). Por el contrario, el RL cognitivo, como el que miden tests clásicos de perseverancia como la WCST, depende probablemente de la DLPFC (Dias et al, 1996). Observaciones anecdóticas (Rolls et al, 1994) y sistemáticas (Bechara et al, 1997) indican que el fallo en el RL afectivo no impide a los lesionados ventromediales adquirir un conocimiento declarativo correcto sobre las contingencias entre estímulos y reforzadores, aunque ese conocimiento no se manifiesta en su conducta. Esto es un indicio más de que el conocimiento declarativo y el conocimiento afectivo (en cuya actualización interviene la VMPFC) se almacenan por separado, y la conducta responde más a este último que a aquél.

El hecho de que los lesionados ventromediales puedan mostrar un conocimiento declarativo correcto de cuál es la opción preferible en una toma de decisiones, al tiempo que escogen alegremente la peor opción, concuerda con el hecho de que la VMPFC no parece intervenir en la toma de decisiones "en frío", es decir, cuando el sujeto sabe que las decisiones no se van a llevar a efecto (Leland & Grafman, 2005). En ese caso las personas lesionadas y las no lesionadas deciden igual de bien, probablemente porque la decisión no les afecta personalmente.

Es muy probable que la VMPFC intervenga en la toma de decisiones que implican riesgo. Sin embargo, no está claro el efecto de la lesión de ésa área sobre la propensión al riesgo. Hay estudios que han observado incremento en la propensión al riesgo, pero en ellos la especificidad de la lesión no se controló (Mavaddat et al, 2000; Rahman et al, 1999). Otros, por el contrario, encontraron reducción (Rogers et al, 1999a). En otros la lesión no parecía tener efecto (Manes et al, 2002), excepto si era extensa (si abarcaba más de un área de entre VMPFC, DLPFC y DMPFC), caso en el que había incremento de la propensión al riesgo. Es posible que la tendencia a optar por las alternativas de riesgo dependa de la extensión de la lesión de una forma compleja (Baxter & Murray, 2001). En personas sin lesión, los estudios con neuroimagen tampoco sugieren una idea clara del papel de la VMPFC en la propensión al riesgo. Algún estudio ha registrado mayor activación de la VMPFC al escoger la alternativa de mayor riesgo que al escoger la más conservadora en un dilema (Ernst et al, 2004). Sin embargo, otro estudio no ha observado tal diferencia (Dickhaut et al, 2003), aunque sí registró una activación ventromedial mayor cuando ambas alternativas tenían valor esperado positivo que cuando lo tenían negativo, sólo en el caso de que ambas alternativas conllevaran riesgo. En animales parece que la lesión ventromedial reduce la propensión al riesgo (Mobini et al, 2002). En la gran mayoría de

los estudios mencionados, los sujetos recibían información sobre el resultado de sus decisiones durante la propia prueba, lo que mezcla probablemente los efectos del aprendizaje con los de la ejecución. La conclusión que puede extraerse de todo esto es que probablemente la VMPFC desempeña algún papel en la resolución del conflicto entre alternativas de diferente riesgo, pero probablemente también el papel de la VMPFC no se limita simplemente a reducir la preferencia por la opción más arriesgada.

Respecto de la impulsividad, los estudios con animales sugieren que la lesión ventromedial incrementa la preferencia por el refuerzo inmediato frente al demorado (Mobini et al, 2002). Aunque en algún estudio se ha obtenido el resultado opuesto (Winstanley et al, 2004), es probable que la contradicción aparente pueda explicarse como efecto del RL. Los estudios mediante neuroimagen con humanos no lesionados sugieren que la expectativa de refuerzo demorado depende de un bucle cortico-basal que interconecta la DLPFC y el estriado dorsal, mientras que la expectativa de refuerzo inmediato se representa en otro bucle que interconecta la VMPFC y el estriado ventral (McClure et al, 2004; Tanaka et al, 2004). La lesión ventromedial parece incrementar la impulsividad en la preparación, y acelera el tiempo subjetivo (lo que posiblemente reduce la capacidad para soportar la demora de un reforzador). No obstante, estos datos fueron obtenidos con medidas indirectas de la preferencia por un reforzador demorado. Son necesarios estudios que analicen el efecto de la lesión ventromedial sobre la impulsividad de forma directa mediante dilemas. Es poco probable, en cualquier caso, que la función de la VMPFC en el control de la impulsividad se limite a reducir la preferencia por los reforzadores menores e inmediatos frente a los mayores demorados. Seguramente, como en el caso de la preferencia por el riesgo, la corteza ventromedial desempeñe un papel más complejo y más difícil de caracterizar.

Quizá la idea que mejor puede resumir el conjunto de datos disponible es que la VMPFC se encarga de representar la expectativa de refuerzo. Para ser completa, la representación de un reforzador esperado debe incluir tanto la demora del reforzador como la varianza de su magnitud, en el caso de que el reforzador cambie de valor de unas ocasiones a otras. Para ser útil, la expectativa de refuerzo debe actualizarse de manera flexible a partir de la experiencia y en función del contexto. Por esta razón, si es cierto que la representación de las expectativas de refuerzo depende de la VMPFC, ésta área cerebral debe estar implicada tanto en la decisión entre alternativas de diferente varianza, como en la decisión entre alternativas con diferente demora del reforzador y en el proceso de actualización de la representación del valor de cada alternativa. Los detalles del funcionamiento de esta área, sin embargo, están por descubrir. También está por aclarar si la incapacidad para llevar a cabo esta función básica de representación del valor puede explicar, por sí sola, todo el conjunto de déficits en conductas complejas que se observan en los lesionados ventromediales, como los problemas de planificación o la carencia de empatía.

La toma de decisiones emocionales es un campo de investigación de posibilidades fascinantes que apenas está comenzando a ser explorado. La toma de decisiones es el proceso en el que la información percibida y la recordada por un organismo se integran para gobernar la conducta, el nexo

de unión entre la percepción, la memoria, y la ejecución motora. Comprender el mecanismo por el que el cerebro es capaz de tomar decisiones sobre los asuntos que afectan al bienestar del organismo nos acercará mucho al entendimiento global de qué somos, y al diseño de tratamientos efectivos para numerosos trastornos que producen enorme sufrimiento a quienes los padecen y a las personas que les rodean. Sin embargo, no cabe duda de que el análisis de un campo tan complejo requerirá del desarrollo de una metodología de complejidad comparable, y no sólo en el aspecto tecnológico, sino también en el conceptual, probablemente mediante el uso de matemáticas avanzadas para la formulación precisa de los modelos. Es difícil creer que la función de la VMPFC en aspectos de la conducta como la propensión al riesgo o la impulsividad se limite a promover su aumento o reducción, sin más matices, y, de hecho, ni siquiera es claro de que esos conceptos sean herramientas especialmente adecuadas para describir la conducta. Aunque son esperanzadoras las aportaciones de algunos grupos de investigación a lo largo de los últimos años, también es patente una tendencia a la inercia en muchos ámbitos, una reticencia a buscar soluciones creativas, una propensión a emplear una y otra vez los mismos paradigmas de investigación con pequeñas variantes. El despegue definitivo de la investigación en este campo dependerá de la capacidad de la comunidad investigadora para llevar a cabo una revolución en la forma de abordar los problemas, para promover la innovación metodológica y aceptar con interés el trabajo creativo.

Agradecimientos: Esta investigación ha sido financiada en parte por el proyecto BSO2003-03723 concedido por el Ministerio de Educación a Antonio Cándido, y realizada bajo los auspicios de una beca de formación de la Junta de Andalucía concedida a David Contreras.

Referencias

Barratt, E. S. (1983). The biological basis of impulsiveness: The significance of timing and rhythm disorders. *Personality and Individual Differences, 4*, 387–391

Baxter, M. G., y Murray, E. A. (2001). Opposite relationship of hippocampal and rhinal cortex damage to delayed nonmatching-to-sample deficits in monkeys. *Hippocampus, 11,* 61–71.

Bechara, A., y Damasio, A.R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior, 52,* 336–372

Bechara, A., Damasio, A.R., Damasio, H., y Anderson, S.W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50,* 7-15.

Bechara, A., Damasio, H., y Damasio, A.R. (2000a). Emotion, decision-making and the orbitofrontal cortex. *Cerebral Cortex, 10*, 295-307.

Bechara, A., Damasio, H., Tranel, D., y Damasio, A.R. (1997). Deciding advantageously before knowing the advantage. *Science, 275*, 1293-1295.

Bechara A., Damasio, H., Tranel, D. y Damasio, A.R. (2005). The Iowa Gambling Task and the somatic marker hipothesis: Some questions and answers. *TCS, 9*, 159-162.

Bechara, A., Tranel, D. y Damasio, H. (2000b). Characterization of the decision-making deficit of patients with ventromedial frontal cortex. *Brain, 123*, 2189-2202.

Bechara, A., Tranel, D., Damasio, H., y Damasio A.R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex, 6*, 215-25.

Berlin, H.A., Rolls, E.T. y Kischka, U. (2004). Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal lesions. *Brain, 127*, 1108-1126.

Blair, R.J.R. (2004). The roles of orbitofrontal cortex in the modulation of antisocial behavior. *Brain and Cognition, 55*, 198-208.

Bowman, C.H. y Turnbull, O.H. (2004). Emotion-based learning on a simplified card game: The Iowa and Bangor Gambling Tasks. *Cognition, 55*, 277-282.

Brand, M., Kalbe, E., Labudda, K., Fujiwara, E., Kessler, J. y Markowitsch, H.J. (2005). Decision-making impairments in patients with pathological gambling. *Psychiatry Research, 133*, 91-99.

Breiter, H., Aharon, I., Kahneman, D., Dale, A. y Shizgal, P. (2001). Functional imaging of the neural responses to expectancy and experience for monetary gains and losses. *Neuron, 30*, 619–639.

Busemeyer, J.R. y Stout, J.C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara gambling task. *Psychological Assessment, 14*, 253- 262.

Chamberlain, S.R., Blackwell, A.D., Fineberg, N.A., Robbins, T.W. y Sahakian BJ (2005). The neuropsychology of obsessive compulsive disorder: The importance of failures in cognitive and behavioural inhibition as candidate endophenotypic markers. *Neuroscience and Biobehavioral Reviews, 29*, 399-419.

Clark, L., Manes, F., Antoun, N., Sahakian, B.J. y Robbins T.W. (2003). The contributions of lesion laterality and lesion volume to decision-making impairment following frontal lobe damage. *Neuropsychologia, 41*, 1474-1483.

Critchley, H., Mathias, C. y Dolan, R. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron, 29*, 537–545.

Damasio, A.R. (1996). *El error de Descartes: La emoción, la razón y el cerebro humano*. Barcelona: Crítica.

Damasio, H., Grabowski, T., Frank, R., Galaburda, A.M. y Damasio, A.R. (1994). The return of Phineas Gage: Clues about the brain from the skull of a famous patient. *Science, 264*, 1102-5. Erratum en: *Science* (1994) *265*, 1159.

Damasio, A.R., Tranel, D. y Damasio, H. (1990). Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behavioral Brain Research, 41*, 81-94.

Dias, R., Robbins, T.W. y Roberts, A.C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature, 380*, 69-72.

Dias, R., Robbins, T.W. y Roberts, A.C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from “on-line” processing. *The Journal of Neuroscience, 17*, 9285–9297.

Dickhaut, J., McCabe, K., Nagode, J.C., Rustichini, A., Smith, K. y Pardo, J.V. (2003). The impact of the certainty context on the process of choice. *Proceedings of the National Academy of Sciences of the USA, 100*, 3536-3541.

Dimitrov, M., Phipps, M., Zahn, T.P. y Grafman, J. (1999). A thoroughly modern Gage. *Neurocase, 5*, 345-354.

Dunn, B.D., Dalgleish, T., y Lawrence, A.D. (2006). The somatic marker hypothesis: a critical evaluation. *Neuroscience and Biobehavioral Reviews, 30*, 239-271.

Ernst, M., Nelson, E.E., McClure, E.B., Monk, C.S., Munson, S., Eshel, N. et al. (2004). Choice selection and reward anticipation: an fMRI study. *Neuropsychologia, 42*, 1585-1597.

Eslinger, P.J. y Damasio, A.R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology, 35*, 1731-41.

Evenden, J.L., (1999). Varieties of impulsivity. *Psychopharmacology,* 146, 348–361.

Fellows, L.K. (2004). The cognitive neuroscience of human decision making: A review and conceptual framework. *Behavioral and Cognitive Neuroscience Reviews, 3*, 159-172.

Fellows, L.K., y Farah, M.J. (2003). Ventromedial frontal cortex mediates affective shifting in humans: Evidence from a reversal learning paradigm. *Brain 126*, 1830-1837.

Fellows, L.K., y Farah, M.J. (2005). Different underlying impairments indecision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cerebral Cortex, 15*, 58-63.

Grafman, J., Schwab, K., Warden, D. y Pridgen, A. (1996). Frontal lobe injuries, violence and aggresion: A report from the Vietnam head injury study. *Neurology, 46*, 1231-8.

Haldane, M. y Frangou, S. (2004). New insights help define the pathophysiology of bipolar affective disorder: neuroimaging and neuropathology findings. *Progress in Neuropsychopharmacological Biology and Psychiatry, 28*, 943-60.

Harlow, J.M. (1868). Recovery from the passage of an iron bar through the head. *Publications of the Massachussets Medical Society, 2,* 327-347.

Holland, P.C. y Gallagher, M. (2004). Amygdala-frontal interactions and reward expectancy. *Current Opinion in Neurobiology, 14*, 148–155.

Kagan, J. (1966). Reflection-impulsivity: The generality and dynamics of conceptual tempo. *Journal of Abnormal Psychology, 71*, 17–24

Kalivas, P.W. y Volkow, N.D. (2005). The neural basis of addiction: A pathology of motivation and choice. *American Journal of Psychiatry, 162*, 1403-1413.

Kahneman, D. (2003). A perspective on judgement and choice: Mapping bounded rationality. *American Psychologist, 58*, 697-720.

Krawczyk, D.C. (2002). Contributions of the prefrontal cortex to the neural basis of human decision making. *Neuroscience and Biobehavioral Reviews, 26*, 631-664.

Kringlebach, M.L. y Rolls, E.T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology. *Progress in Neurobiology, 72*, 341-372.

Lee, J.S., Kim, B.N., Kang, E., Lee, D.S., Kim, Y.K., Chung, J-K, et al. (2005) Regional cerebral blood flow in children with attention deficit hyperactivity disorder: comparison before and after methylphenidate treatment. *Human Brain Mapping, 24*, 157-164.

Leland, J.W. y Grafman, J. (2005). Experimental tests of the somatic marker hypothesis. *Games and Economic Behavior, 52*, 386-409.

Loewenstein, G.F, Weber, E.U., Hsee, C.K. y Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127*, 267-286.

Maia, T.V. y McClelland, J.L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences of the USA, 101*, 16075-16080.

Malloy, P. y Grace, J. (2005). A review of rating scales for measuring behavior change due to frontal systems damage. *Cognitive and Behavioral Neurology, 18*, 18

Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M. et al. (2002). Decision-making processes following damage to prefrontal cortex. *Brain, 125*, 624-639.

Mavaddat, N., Kirkpatrick, P.J., Rogers, R.D. y Sahakian, B.J. (2000). Deficits in decision making in patients with aneurysms of the anterior communicating artery. *Brain 123*, 2109-2117.

McAlonan, K. y Brown, V.J. (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behavioural Brain Research 146*),97-103

McClure, S.M., Laibson, D.I., Loewenstein, G. y Cohen, J.D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306*, 503-507

Mobini, S., Body, S., Ho, M.Y., Bradshaw, C.M., Szabadi, E., Deakin, J.F.W. et al. (2002). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology, 160*, 290-298.

O'Driscoll, K. y Leach, J.P. (1998). "No longer Gage": An iron bar through the head. *British Medical Journal, 317*, 1673-4.

Petrides, M. (2005). Lateral prefrontal cortex: architectonic and functional organization. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 360*, 781-795.

Phelps, E.A, y LeDoux, J.E. (2005). Contributions of the amygdala to emotional processing: from animal models to human behavior. *Neuron, 48*, 175-187.

Rahman, S., Sahakian, B.J., Hodges, J.R., Rogers, R.D. y Robbins, T.W. (1999). Specific cognitive deficits in mild frontal variant frontotemporal dementia. *Brain, 122*, 1469-1493.

Robinson, T. y Berritge, K. (2003). Addiction. *Annual Review of Psychology , 54,* 25-53.

Rogers, R., Everitt, B.J., Baldacchino, A., Blackshaw, A.J., Swainson, R., Wynne, K. et al. (1999a). Dissociable deficits in the decision making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to the prefrontal cortex, and tryptophan depleted normal volunteers: Evidence for monoaminergic mechanisms. *Neuropsychopharmacology, 20*, 322–39.

Rogers, R.D, Owen, A.M., Middleton, H.C., Williams, E.J., Pickard, J.D., Sahakian, B.J. et al. (1999b). Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *The Journal of Neuroscience, 20*, 9029-9038.

Rolls, E.T. (2000). The orbitofrontal cortex and reward. *Cerebral cortex 10*, 284-294.

Rolls, E.T., Hornak, J., Wade, D., y McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery, and Psychiatry, 57*, 1518-1524.

Rustichini, A., Dickhaut, J., Ghirardato, P., Smith, K. y Pardo, J.V. (2005). A brain imaging study of the choice procedure. *Games and Economic Behavior 52*, 257-282

Sabbagh, M.A. (2004). Understanding orbitofrontal contributions to theory-of-mind reasoning: implications for autism. *Brain and Cognition, 55*, 209-219.

Séguin, J.R. (2004). Neurocognitive elements of antisocial behavior: relevance of an orbitofrontal cortex account. *Brain and Cognition, 55*, 185-197.

Stout, J.C., Rodawalt, W.C. y Siemers, E.R. (2001). Risky decision making in Huntington's disease. *Journal of the International Neuropsychological Society, 7*, 92-101.

Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamoto, Y. y Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature Neuroscience, 7*, 887-893.

The Mysterious Orbitofrontal Cortex (2000). *Cerebral Cortex, 10*.

Tomb, I., Hauser, M., Deldin, P., y Caramazza, A. (2002). Do somatic markers mediate decision on the gambling task? *Nature Neuroscience, 5*, 1103-1104.

Turnbull, O.H., Evans, C.E.Y., Bunce, A., Carzolio, B. y O'Connor, J. (2005). Emotion-based learning and central executive resources: An investigation of intuition and the Iowa Gambling Task. *Brain and Cognition, 57*, 244-247.

Winstanley, C.A., Theobald, D.E.H., Cardinal, R.N. y Robbins, T.W. (2004). Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *The Journal of Neuroscience, 24*, 4718-4722.

CAPÍTULO 2

Facilitation and interference of behavioral responses by task-irrelevant emotional stimuli

Facilitation and Interference of Behavioral Responses by Task-irrelevant Affect-laden Stimuli

David Contreras[1], Alberto Megías, Antonio Maldonado, Antonio Cándido, and Andrés Catena

Universidad de Granada[2]

---

1 Correspondence concerning this article should be addressed to David Contreras, Departamento de Psicología Experimental, Universidad de Granada, Spain. E-mail: davcr@ugr.es

2 This work was supported by the Junta de Andalucía (P06-HUM02375) and the Spanish Ministry of Science and Innovation (SEJ2006-11906/PSIC) grants to Andrés Catena and Antonio Maldonado, respectively.

Published as:

Contreras, D., Megías, A., Maldonado, A., Cándido, A., & Catena, A. (2012). Facilitation and interference of behavioral responses by task-irrelevant affect-laden stimuli. *Motivation and Emotion*. doi:10.1007/s11031-012-9327-0

Abstract

Emotional interference on behavior is commonly observed when task-irrelevant negative stimuli appear before behavioral targets. One explanation postulates that affect-laden stimuli readily capture attention, interfering with the processing of the upcoming target. Emotional stimuli might also preactivate motor programs incompatible with the demanded response. Using a cued go/no-go procedure we showed that task-irrelevant unpleasant stimuli cause interference or facilitation depending on their onset asynchrony relative to the target. We observed interference with short (200 ms) stimulus-target asynchronies and facilitation for longer ones (600 ms), both for key press (Experiment 1) and key release (Experiment 2) responses. The interference effect is compatible with an attentional explanation, but the behavioral facilitation is hard to accomodate within either attentional or motor accounts. This interference-facilitation pattern can be explained assuming that once the attentional effect subsides, emotional processing may enhance the perceptual processing of the stimuli, or lower the decision threshold, thereby facilitating the response selection process.

Keywords: emotion, attention, motor, perception, decision-making

Facilitation and Interference of Behavioral Responses by Task-irrelevant Affect-laden Stimuli

The mutual influences between cognitive and emotional processing have been the subject of a recent surge of scientific interest (Dolan, 2002; Zald, Matson, & Pardo, 2002). Studies on the emotional modulation of cognitive processes frequently adapt classical behavioral tasks by including some sort of emotional manipulation. Usually, the targets in the behavioral task are simply replaced by affect-laden stimuli (words or images of positive or negative valence). Many experimental studies, thus, introduce emotion just as the target of discrimination in a discrimination task: e.g., the participant has to decide whether a word shown on a screen has negative or positive affective value. In those experiments, there is no explicit aim to change the emotional state of the participant, i.e., to make the participant *feel* anything. In other words, the emotional value is treated rather as an attribute of the discriminandum than as a state of the subject. In such circumstances, it is hard to tell whether the effects observed are attributable to emotional processing or to the fact that the discrimination task is a peculiar one (Goldstein et al., 2007). Only a few studies have addressed the question of how emotional content affects behavior when it is incidental to the task at hand. To make emotional processing incidental to the task, some of them have instructed participants to discriminate low level features (such as typeface) of emotion-laden words (Goldstein et al., 2007). Others have showed the participants task-irrelevant emotional stimuli (Pereira et al., 2004; Pereira et al., 2006; Phelps, Ling, & Carrasco, 2006; Bocanegra & Zeelenberg, 2009; Becker, 2009). The second strategy mimics circumstances encountered in a great number of situations in everyday life (i.e.: driving

after having watched a road accident) (Megías, Maldonado, Catena, Di Stasi, Serrano, & Cándido, 2011; Megías, Maldonado, Cándido, & Catena 2011).

Using an implicit emotional task, Pereira et al. (2004) found that visual detection times were slower after watching negatively-valenced task-irrelevant images than after positive or neutral ones (the targets themselves being emotionally neutral). They attributed this effect to the interference caused by the motor program triggered by the emotional stimuli (avoidance response for negative pictures) on the response demanded by the task (keystroke, that they considered an approach response). Using a similar procedure, Pereira et al. (2006) have observed both transient and sustained emotional modulation of visual detection of targets after watching task-irrelevant mutilation pictures. The sustained effect lingered several seconds after image offset, and was apparent when the pictures were blocked according to their emotional value. Conversely, when the emotional value of the pictures was randomized in the temporal sequence, only transient effects –affecting just the first trial after image offset– were observed. The authors believed that the former effect depended on the link between the induced emotional state and approach/withdrawal processes, so that watching mutilation images engaged the subjects in avoidance behaviors incompatible with the approach response demanded by the task (keystroke). An alternative explanation for the sustained effect suggested in that study was based on the idea that affect-laden stimuli may produce different degrees of freezing. A freezing response is incompatible with any other response, and, therefore, the same size of effect should be observed either for approach or withdrawal responses. For the transient effect, however, these authors favored an attentional account: unpleasant pictures reduced the available processing resources as they have a greater capability for capturing attention.

Buodo, Sarlo, and Palomba (2002) found transient emotional interference effects in an auditory discrimination task. In their experiment, a task-irrelevant image was shown in each trial for 6 seconds, with its onset 1 or 4 seconds before the target sound was played. They observed that erotic and mutilation images caused an increase in discrimination reaction times when the pictures appeared 1 second before the target, but not when they appeared 4 seconds before. Similarly, using a rapid serial visual presentation procedure, Most, Chun, Widders, and Zald (2005) have also observed that target detection accuracy was lower after emotional pictures relative to non-emotional ones when the SOA was 200 ms, but there were no differences when the SOA was 800 ms.

These remarkable results show that the way task-irrelevant emotional stimuli interfere with the response to the target may depend on the picture-target onset asynchrony (SOA). As mentioned, emotional interference has been attributed to the allocation of attentional resources to the emotional stimuli or, alternatively, to the preactivation of motor processes by them. An attentional account predicts that the emotional interference will fade shortly after attention is disengaged from the emotional stimuli. Since the attentional dwell time estimated from visual search experiments is believed to be around 250 ms (Theeuwes, Godijn, & Pratt, 2004), it is reasonable to expect that attention should be away from the emotional stimuli at around 500-800 ms (Pereira et al., 2006), even if it is assumed that disengaging attention from emotional stimuli takes longer than disengaging it from neutral ones. On the other hand, the motor hypothesis predicts that response facilitation or response interference will be observed depending on whether the motor response demanded by the task (e.g.: keystroke) is compatible or incompatible with the motor response presumably preactivated by the emotional stimulus (e.g.: withdrawal). An integration of these two accounts predicts

maximal interference when the demanded response is incompatible with the one preactivated by a task-irrelevant emotional stimulus and the emotional stimulus is shown shortly before the target (attention is still engaged in the emotional stimuli and the preactivated motor response must be inhibited). If the response demanded is incompatible with the emotional stimulus valence but the emotional stimulus is shown long before the target, only motor interference should occur, and the observed effect should be weaker.

In our study, we used a cued emotional go/no-go task to specifically investigate the emotional modulation of behavioral responses by the visualization of task-irrelevant emotional stimuli shown at different SOAs before the target. Originally introduced by Donders (1868/1969), the go/no-go task requires the participant to discriminate between two types of stimuli presented in random sequence –go and no-go targets– and respond only to the former type. In a typical experiment, go targets are more frequent than no-go ones, and this is assumed to create a predisposition to respond in every trial, which has to be overcome in the rare cases when a no-go target appears. When go and no-go emotionally charged stimuli are used, it is possible to measure the modulation of performance by emotional processing. Differences between the reaction times for neutral and emotional go targets have been interpreted as reflecting an emotional bias on the tendency to approach or avoid the target (Hare, Tottenham, Davidson, Glover, & Casey, 2005). Differences in error rates for emotional and neutral no-go targets are commonly interpreted as the result of the emotional modulation of the ability to inhibit motor responses (Schultz, Fan, Magidina, Marks, Hahn, & Halperin, 2006). In our experiments, targets were, however, emotionally neutral. The emotionally-charged stimuli were instead task-irrelevant pictures shown before target onset.

In a cued go/no-go procedure, a warning signal indicates the probable identity of the upcoming target. Anticipatory activity throughout the foreperiod –the interval between the warning cue and the target– leads to faster responses to the target (Niemi & Näätänen, 1981). It has been shown that warning signals trigger motor activation (as measured by electromyography) in an automatic way and that the probability of commission errors depends on the strength of this activation (Boulinguez, Jaffard, Granjon & Benraiss, 2008, see also Nobre, Correa & Coull, 2007). In our experiments, a predictive cue was shown before the emotional picture in each trial. Thus, two forces could contribute independently to create a proneness to respond in our experiments: the predictive cue, and –according to the motor preactivation hypothesis– the emotional picture, depending on its valence. We also manipulated the onset asynchrony between the emotional image and the target that followed it, providing enough time for the disengagement of attention from the emotional stimuli in some experimental conditions but not in others. Moreover, we manipulated the cue-target onset asynchrony, to explore the effect on the response of the interval between the cue and the task-irrelevant emotional stimuli. Lastly, the response demands were manipulated between experiments, requiring from the participant two types of responses: key strokes (Experiment 1) and key releases (Experiment 2). If visualizing emotional pictures causes the activation of specific motor programs, a change in the demanded response might alter or even reverse the effect of the emotional pictures on reaction time. We also expected a tendency towards interference when the emotional picture was shown shortly before the target (for delays within the attentional dwell time), but not when this interval was long enough to allow attention to disengage from the picture.

Experiment 1

In this experiment we investigated the effect of the incidental visualization of emotional pictures on the speed and accuracy of keystroke responses in a cued go/no-go task. A bias to respond or withhold motor responses was elicited by a highly predictive cue, which was displayed 800 ms or 400 ms before the target. Task-irrelevant emotional distractors (pictures) were presented at two different time lags (200 ms or 600 ms before the target) within the cue-target interval in order to affect the ongoing mental activity (attention engagement and motor pre-activation) elicited by the cue. The shorter picture-target interval (200 ms) ocurred for trials with cue-target asynchrony of 400 or 800 ms whereas the longer picture-target interval (600 ms) only happened in trials with 800 ms cue-target asynchrony. A no-picture condition served as the baseline for measuring the interference produced by the distractor pictures.

Method

Participants. Twenty (2 male) undergraduate students at the Faculty of Psychology of the University of Granada (Spain) ranging in age from 18 to 25 (mean 20.8) took part in this study, in exchange for course credit.

Apparatus and stimuli. The stimuli were displayed on a 15-inch LCD monitor, at 1024x768 pixels resolution, 32-bit color depth. Refresh rate was 60 Hz (6 frames ≈ 100 ms). The task was coded in Java using PXLab library for psychophysical experiments (Irtel, 2006). The predictive cue was a white bar, either vertical (100 x 300 pixels, which was predictive of a go target) or horizontal (300 x 100 pixels, predicting a no-go target),

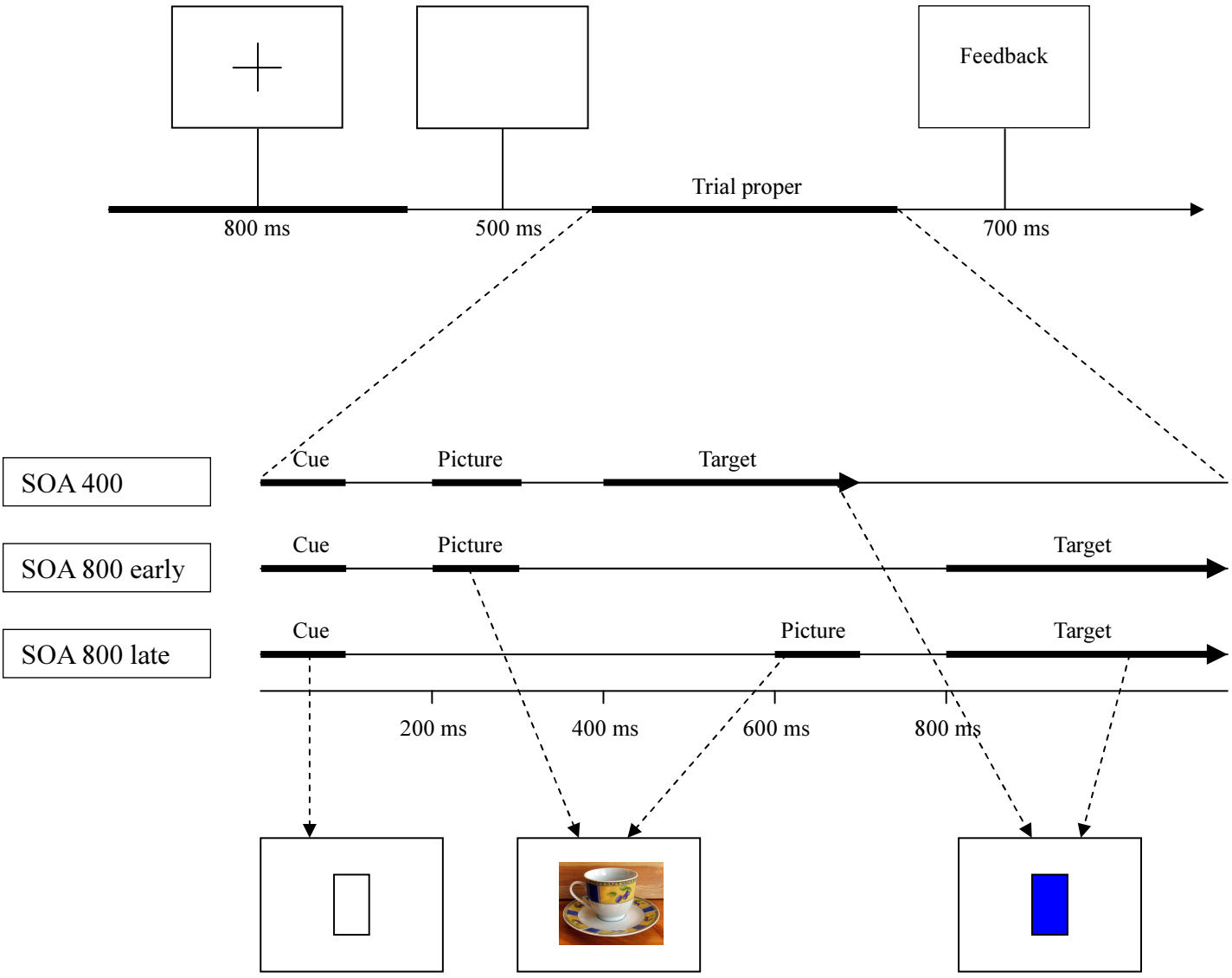

Figure 1. Sequence of events in typical go trials in Experiments 1 (space bar key press) and 2 (space bar release).

shown at the screen center. The target was a bar whose size, location, and orientation were identical to those of the cue presented before, but colored in blue (go target) or green (no-go target). Affective pictures belonging to four categories were selected from the International Affective Picture System (IAPS) (Lang, Bradley, & Cuthbert, 2005): mutilation (average valence: 1.32, max=1.42, min=1.18, average arousal: 7.55), babies and children (average valence: 8.34, max=8.43, min=8.18, average arousal: 3.76),

neutral common objects (a towel, a stool, a cabinet, a plate, average valence = 5.25, max = 5.69, min = 4.7, average arousal = 2.43), and erotic (couples, average valence = 7.21, max = 7.42, min = 7.06, average arousal = 6.67). Four images were selected for each of the four categories. Valence and arousal values are taken from the Spanish female population tables (Moltó, Montañés, Poy, Segarra, et al., 1999). The original IAPS pictures were scaled down so that their height was 384 pixels or their width was 512 pixels, whichever was more restrictive, but maintaining their original aspect ratio. All visual stimuli were shown over a black background. Participant responses were recorded using a standard PC keyboard.

Design and procedure. All participants gave written consent to their participation before the beginning of the experiment, after being warned that the task involved briefly watching some images that could be deemed offensive, and being also reminded of their right to withdraw their participation without giving up any of the benefits derived from it. Then the instructions asked them to press the space bar as fast as they could whenever they saw a blue bar and to refrain to respond when they saw a green bar.

Each trial started with a fixation point (a 50 x 50 pixel cross on the screen center) that lasted for 800 ms, immediately followed by a blank screen for 500 ms (see Figure 1). The predictive cue (either a vertical or a horizontal white bar) was then shown for 100 ms. Half of the trials had a go (vertical) cue, and the other half a no-go (horizontal) cue. Go cues were followed by a go target (blue) in most cases (80%), and by no-go targets in the rest of the trials (20%). Likewise, a no-go cue was followed by a no-go target in 80% of the occasions, and a go target in the remaining 20%. The stimulus onset asynchrony (SOA) between cue and target was 400 ms in one third of the

trials and 800 ms in the rest. The target remained on-screen for a maximum of 1000 ms or until the participant responded. After the offset of the target a message was shown on the center of the screen for 700 ms with the word "*Correcto*" (the Spanish word for "correct") and, for go trials, the reaction time in ms below it, or the word "*Fallaste*!" ("You failed!").

On 4 out of every 5 trials an affective image from IAPS (Lang et al., 2005) was shown between cue and target for 100 ms. The image belonged to one of the four aforementioned categories (mutilation, babies, categories, neutral). On the remaining trials, no image was shown. We manipulated the time onset of the image in order to explore the temporal course of the affective interference. When the cue-target SOA was short (SOA400 condition), the target onset was 200 ms after the image onset. However, when the cue-target SOA was long, the image onset could be delayed either 200 ms (SOA800-early) or 600 ms (SOA800-late) from cue onset.

There were 5 affective (mutilation, babies, erotic, neutral and no-picture) x 3 timing (SOA400, SOA800-early, SOA800-late) conditions. Each affective x timing condition comprised 40 trials, 20 with go target and 20 with no-go target. On 80% of the trials, the cue was congruent with the target. Therefore, of the 40 trials in each condition, 16 were congruent go trials, 16 were congruent no-go trials, 4 were incongruent go trials, and the remaining 4 were incongruent no-go trials. As there were only 4 incongruent go and 4 incongruent no-go trials in each affective x timing condition, each affective picture category (babies, mutilations, erotic, neutral) contained just four pictures, so that every picture was shown at least once in each kind of trial. This meant that each image was shown 10 times in each affective x timing condition, 8 of them in congruent trials (four go, four no-go), and 2 in incongruent trials (one go, one

no-go), totalling 30 repetitions in the whole task (ten in each of the three timing conditions with pictures, namely SOA400, SOA800-early, SOA800-late). Trials were blocked according to timing condition, giving rise to three blocks for conditions with task-irrelevant image (SOA800-early, SOA800-late and SOA400), and two blocks for conditions without picture, one with long (800 ms) and another with short (400 ms) SOA. Within each block, the order of the trials and images was randomized for each participant. The blocks without pictures had a total of 40 trials. The blocks with pictures contained 40 trials of each affective condition, totalling 160 trials per block. The total number of trials was 560.

The experiment was carried out according to one of the following sequences of blocks: 1) SOA400 without pictures, SOA400 with pictures, SOA800 without pictures, SOA800-late (or SOA800-early), SOA800-early (or SOA800-late); 2) SOA800 without pictures, SOA800-late (or SOA800-early), SOA800-early (or SOA800-late), SOA400 without pictures, SOA400 with pictures. Blocks of trials with the same SOA were always placed together, and blocks without pictures came always before blocks with pictures. These sequences were chosen so that the sequential effects caused by changes of SOA were minimal in the blocks with images, since they always came after one block with identical SOA. The sequence for each participant was randomly selected.

The interval between the participant response and the beginning of the following trial was adjusted to compensate for the variations in RT, so that a constant interval of 2400 ms elapsed between the onset of the target on one trial and the appearance of the fixation cross on the following trial. In this way the whole trial lasted always either 4.1 or 4.5 seconds (depending on whether the cue-target SOA was 400 ms or 800 ms).

There was a 1-minute break after each two consecutive blocks, making the duration of the whole experiment about 50 minutes.

Data Analysis

Median RTs for correct go trials were computed for each participant and condition. Commission error rates were computed for no-go trials in each condition. Averages across subjects of the medians of reactions times and error percentages are displayed in Table 1. Data analysis was organized in two sets of repeated-measures

	Experiment 1			Experiment 2		
	SOA800 late	SOA800 early	SOA 400	SOA800 late	SOA800 early	SOA 400
Babies	335 (9.00)	310 (4.50)	318 (14.25)	364 (8.75)	336 (8.44)	356 (12.34)
Erotic	349 (12.00)	311 (5.50)	327 (13.25)	384 (11.41)	338 (3.75)	349 (14.84)
No Picture	319 (6.00)	319 (6.00)	301 (9.25)	337 (10.62)	337 (10.62)	326 (15.31)
Neutral	347 (8.00)	316 (4.00)	323 (12.75)	362 (12.66)	346 (6.25)	352 (15.78)
Mutilations	354 (13.00)	306 (5.50)	338 (11.50)	377 (17.81)	330 (12.19)	361 (16.87)

Table 1: Averages of median reaction times (in milliseconds) and error percentages (within parenthesis) for each timing condition and picture contents.

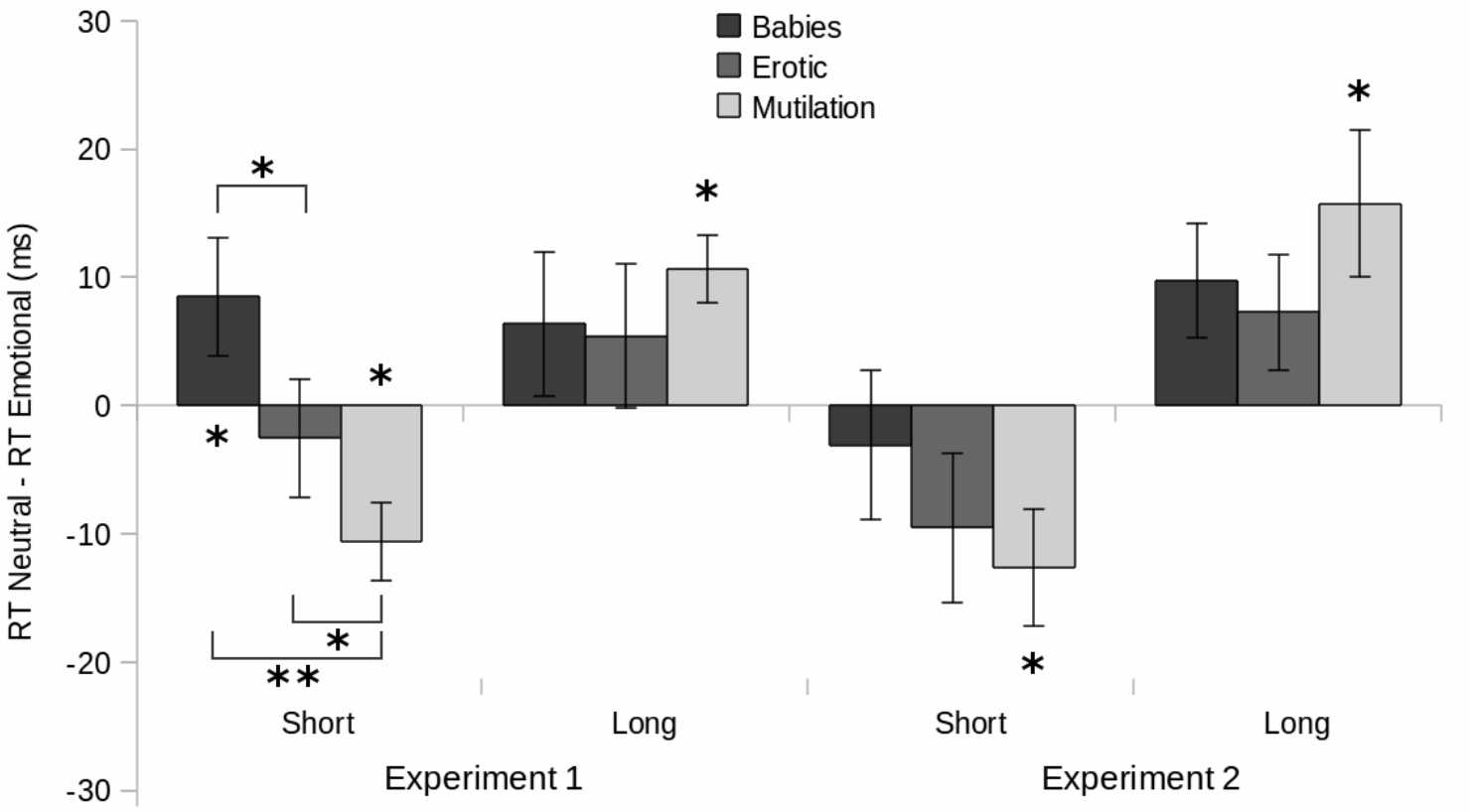

Figure 2. Interference caused by pictures of different affective value for each picture-target asynchrony (Short: 200 ms, Long: 600 ms). Vertical dimension represents the difference between the RT in the neutral condition and the RT in each affective condition. Positive values indicate interference and negative ones facilitation. Vertical bars represent the standard error of the difference. The left panel represents the results of Experiment 1 (demanded response: key press), and the right panel, Experiment 2 (demanded response: key release). The pattern of interference and facilitation is similar in both experiments. Significant differences are indicated using * for $p < .05$ and ** for $p < .01$. An asterisk above or below of a bar indicates a significant difference between the RTs to that type of pictures and to neutral ones, while over a bracket, it indicates a significant difference between the two conditions connected by the bracket.

ANOVAs. In the first set we explored differences between the conditions without emotional pictures (Baseline analysis), while in the second one we examined the differences between the conditions with pictures (Emotion analysis). In the baseline analysis, for the no-picture conditions, we performed a 2 (SOA: 400 ms and 800 ms) x 2 (Cue-target congruence: congruent and incongruent) repeated-measures ANOVA to explore the differences in reaction time and commission errors. For the Emotion

analysis, median reaction times for the conditions with pictures were submitted to a repeated-measures ANOVA with timing condition (SOA400, SOA800-early and SOA800-late), affective content of the images (babies, erotic, mutilations, and neutral) and cue-target congruence (congruent and incongruent) as within-subject factors. The percentages of commission errors for no-go trials were submitted to a repeated-measures ANOVA with timing condition (SOA800-late, SOA800-early and SOA400), affective content (babies, erotic, neutral, and mutilation), and cue-target congruence (congruent, incongruent) as within-subject factors. To further explore significant effects we first checked the differences between SOA400 and SOA800-late, as they shared the same short picture-target onset asynchrony (200 ms). These two conditions were then pooled together if no differences were observed. Accordingly, we will use the term "short asynchrony" for the pool SOA400-SOA800-late and "long asynchrony" for SOA800-early. Finally, we performed post-hoc LSD comparisons. A significance level of 0.05 was set up for all statistical decisions.

Results and Discussion

Go trials. Omission errors occurred in only 0.1% of the go trials, so no attempt to analyze omission errors was made. The baseline analysis for reaction times revealed significant main effects of cue-target SOA, F (1, 19) = 6.184, p = .022, and cue-target congruence, F (1, 19) = 10.68, p = .004, but no interaction. Responses were slower for the long (800 ms) than for the short (400 ms) SOA (319 ms and 301 ms, respectively). Also, responses were faster for the cue-target-congruent condition than for the incongruent condition (301 ms and 318 ms, respectively), indicating that the participants used the cue for preparing their response to the target.

Emotion analysis (see Figure 2). The ANOVA showed significant main effects of affective content, F (3, 57) = 5.04, p = .003, R^2 = .21, and timing condition, F (2, 38) = 15.91, p < .001, R^2 = .46, and a significant interaction between affective content and timing, F (6, 114) = 3.09, p = .007, R^2 = .14. We also observed a significant effect of congruence, F (1, 19) = 8.07, p = .01, R^2 = .29. No other effects were significant.

SOA800-late and SOA400 conditions showed no interaction with affective content either in median RTs or in errors, both F (3, 57) < 1, so they were pooled together to obtain the short asynchrony reaction times. Post-hoc LSD analyses of the affective content by timing interaction revealed than in the short asynchrony conditions, reaction times for mutilation pictures were slower than those in the remaining conditions, all *p's* < .035, t(19) > 2.28, and reactions after babies pictures were faster than for all other pictures, all p's < .012, t(19) > 2.80, but erotic pictures did not differ from neutral ones, p > .30. In stark contrast, in the long asynchrony condition reaction times for mutilation pictures were faster than for neutral ones, p = .044, t(19) = 2.16, but not different from those for babies or erotic pictures, both p's > .20. Moreover, reaction times for positive pictures (either erotic or babies) did not differ from those for neutral ones (p's > .28) or between them (p = .81). On the other hand, reaction times in the long asynchrony conditions were faster than those in the short asynchrony ones for all emotional contents, all *p's* < .01, t(19) > 2.89. Interestingly, the difference between the short and the long asynchronies was larger for mutilation than for all other pictures, all *p's* < 0.02, t(19) > 2.59.

In summary, relative to the neutral condition mutilation pictures interfered with the response if the picture preceded the target by 200 ms (short asynchrony), but facilitated the same response when the picture came 600 ms before the onset of the

target (long asynchrony). Babies pictures facilitated the response for the short picture-target asynchrony, but caused no effect for the long asynchrony. Erotic pictures did not differ from the neutral ones. These results support the idea that mutilations have qualitatively different effects in these two time lags rather than one transient effect that fades in less than 600 ms. To further characterize those two separate effects, it is important to complement the measures of reaction speed with those of decision errors. Decision errors on go trials cannot be observed directly, and their best available estimate are response error rates on no-go trials. The estimate is reliable if we assume the decision process is symmetric for both types of target (go and no-go), despite the fact that the motor responses are not (Gómez, Ratcliff & Perea, 2007).

No-go trials. Commission errors occurred on 9.2% of no-go trials. No differences were observed in the baseline analysis. Emotion analysis showed only main effects of timing, F (2, 38) = 5.89, p = .006, R^2 = .19. Error rates were higher in the short asynchrony conditions (SOA800-late, 10.50, and SOA400, 12.94) than in the long asynchrony one (SOA800-early, 4.87), which indicates that errors depended on the picture-target interval, but not on the affective content.

These results indicate that mutilation pictures shown near the target onset may interfere with motor responses. Crucially, response facilitation, rather than interference, was observed when mutilation pictures appeared 600 ms before the target. This last result is in apparent contrast with Pereira et al. (2006) results showing that interference effects lasted for 500-700 ms. However given that their pictures were displayed for 500 and 2000 ms, it is difficult to compare both results. At any rate, Pereira et al. (2006) results agree with ours in that, even with longer viewing times, the effect of mutilation pictures fades very fast. Thus, briefly watching mutilation pictures had two opposite

effects on the speed of response, depending on the picture-target asynchrony. The result observed in the short asynchrony conditions is consistent both with the attentional and the motor preactivation hypothesis. However, the facilitation effect for the longer asynchrony is problematic both to the attentional and the motor pre-activation hypotheses since they predict interference or no effect when mutilation pictures are displayed some time (600 ms) before the target. Neither the reduction in attentional resources nor the preactivation of withdrawal response tendencies (as expected for mutilation pictures) explains the speed-up of responses observed in that condition, unless we assume that watching a mutilation picture preactivates the response of pressing a key. Testing this last possibility was the purpose of Experiment 2, in which we asked the participants to release the key instead of pressing it..

In the general discussion we put forward two possible explanations for this late facilitation of responses by mutilation pictures, in terms of perceptual enhancement (Pessoa, 2008; Phelps, Ling, & Carrasco, 2006; Bocanegra & Zeelenberg, 2009; Becker, 2009) or a relaxation of the decision criterion (Simen, Cohen, & Holmes, 2006).

Experiment 2

In this experiment we changed the response demands to further discriminate the motor preactivation and the attentional accounts. Several studies have reported modulations of reaction time attributable to the emotional content of the target stimuli in speeded reaction tasks, that could be reversed by changing the motor response demanded from the participant. For instance, Solarz (1960) asked students to evaluate

affect-laden words as positive or negative by pulling or pushing a lever, and he found they were faster to do the evaluation if the response mapping was pulling-positive and pushing-negative than when the reverse mapping was used. Chen & Bargh (1999) replicated Solarz (1960) results, and extended them to a simple detection task, where participants had to always pull (or always push) a lever as fast as possible whenever a stimulus (word) appeared. They found the participants to be faster to react to positive stimuli if the response demanded was pulling but faster to react to negative ones if the response demanded was pushing, which suggests that the crucial aspect is the congruence between the stimulus valence and the motor action, rather than the correspondence between each action and its meaning for the task (the response mapping).

Duckworth, Bargh, Garcia, and Chaiken (2002) obtained the same results using completely novel visual stimuli -created specifically for the experiment-, showing that the motor effect of pleasant or unpleasant stimulus does not require a previous learning of the association between the stimulus and a motor response. The actual motor response, however, may be selected by the brain by virtue of a previously learned arbitrary association between *valence* and response, and this has been supported by the literature (e.g., McCall, Tipper, Blascovich, & Grafton, 2011).

Wentura, Rothermund, and Bak (2000) found effects analogous to those just described in a task in which the participants were demanded a much simpler response (either pressing or releasing a key): the reaction was faster when subjects had to press a key in response to a positive word, and also when they had to release the key in response to a negative one. In all the studies mentioned so far, the affect-laden stimulus was the target, but, as mentioned in the introduction, other studies have recorded effects

on reaction times of valenced stimuli that were incidental to the experimental task, and at least some of them attributed these effects to the activation of motor programs by those stimuli (Pereira et al., 2004; Pereira et al., 2006). On the other hand, general links between affective states (fear, anger) and action tendencies (freezing, flight, attack) have been widely documented across species in the ethological and neurophysiological literature (Blanchard & Blanchard, 1988; Lang, Bradley & Cuthbert, 1998; LeDoux & Phelps, 2000)

Therefore, at least in principle, even the slowdown of responses observed for the short picture-target asynchrony might be caused by some form of motor interference due to emotional processing. Changing the response demanded in the task may help rule out that possibility: if the effect of mutilation images on reaction times arose just because they primed a particular motor response, we would expect a change in the pattern of results if the response demanded now is antagonistic to the one demanded before. For example, it might happen that reaction times were longer when a mutilation picture appeared 200 ms before the target simply because processing the picture preactivates certain motor programs, which are incompatible with pressing a key. In this case, the effect of image processing on reaction times should be quite the opposite if the participant was required to release the key for go targets, instead of pressing it. The same can be said about the effect of mutilation pictures when they are shown 600 ms before the target: if the decrease in reaction times occurs because, for some reason, mutilation pictures preactivate keystroke responses in those circumstances, replacing the demanded response by its antagonist should cause a reversal in the effect. The attentional account, on the other hand, predicts no change in the pattern of results for a change in the demanded response.

Method

Participants. Twenty undergraduate students (4 male) at the Faculty of Psychology of the University of Granada, ranging in age from 18 to 26 (mean 20.5) took part in this study in exchange for course credit.

Stimuli and apparatus. The stimuli and the apparatus used in this experiment were identical to those in Experiment 1.

Design, procedure and data analyses. After giving written informed consent, the participants read the instructions of the task. The design and the data analyses of this experiment were the same as those of Experiment 1. The only difference was the response demanded for go trials. As in Experiment 1, only those trials with a correct response were included in RT analyses. Participants had to keep the space bar pressed during the whole session and release it only when a go target was displayed. After releasing the bar, they should press it again. If a participant forgot to press the space bar after having released it, the fixation point of the following trial stayed on screen, and the trial did not start until the participant pressed the space bar again. Averaged median RTs and error percentages are displayed in Table 1.

In order to test for differences in the response patterns between Experiments 1 and 2, the RTs of both experiments were submitted to an ANOVA with response demanded (key press vs key release) as between-groups factor and timing, affective content, and cue-target congruence as within-subject factors.

Results and Discussion

Go trials. Omission errors occurred in 0.5% of the go trials. Baseline RT analysis showed only a main effect of congruence, F (1, 19) = 9.78, p = .006, R^2 = 0.34, with faster reaction times for congruent (322 ms) than for incongruent trials (342 ms). Emotion analysis showed significant main effects of congruence, F (1, 19) = 19.58, p < .001, R^2 = 0.51, and timing condition, F (2, 38) = 6.29, p < .004, R^2 = .25, and a significant interaction between affective contents and timing, F (6, 114) = 4.22, p < .001, R^2 = 0.18. No other effects were significant.

As in Experiment 1, SOA800-late and SOA400 conditions showed no interaction with affective content, so we pooled them together in a single Short picture-target asynchrony condition. The a posteriori LSD analysis of the affective content by timing interaction showed that reaction times were slower for mutilation images than for neutral ones in the short asynchrony condition (p = .044, t(19) = 2.16). Again, responses were faster for mutilation than for neutral pictures in the long asynchrony condition (p = .018, t(19) = 2.596) (see Figure 2). On the other hand, differences between short and long asynchronies were significant in mutilation, babies and erotic pictures (*all p's* < .03, t(19) > 2.41), but not in neutral ones (p = 0.259, t(19) = 1.16)

The comparison between the two experiments yielded a marginally significant main effect of response demanded, F (1, 38) = 3.42, p = .072, which can be attributed to a difference in how hard is to perform each of the responses that were demanded. Also significant was the interaction between response demanded and congruence, F (1, 38) = .046. The Picture x SOA interaction was significant, F (6, 228) = 6.11, p < 0.01, as expected. Crucially, however, the Picture x SOA x response effect was not significant (p = .21), which means that the combined effect on RTs of timing and picture content was similar in both experiments. There were no other significant effects.

In summary, relative to mutilation pictures, and despite the change in the required response, the results of Experiment 2 are similar to those of Experiment 1, as responses were slowed down at the short picture-target asynchrony, but they were speeded-up at the long asynchrony. Leaving aside the question of whether pressing/releasing a keyboard key can be considered as real approach/withdrawal responses, these are motor responses with very different motor programs, and therefore, it seems unlikely that the emotional effects on reaction times can be attributed to the preactivation of specific motor responses.

No-go trials. Commission errors occurred on 11.93% of no-go trials. Higher error rates were observed for the short than for the long cue-target interval in the no-pictures conditions. Emotion analysis showed a significant main effect of affective content, F (3, 57) = 3.90, p = .013, R^2 = 0.18, and significant main effect of timing, F (2, 38) = 4.12, p = .024, R^2 = 0.17. Error percentages were higher for mutilation (15.62%) than for the remaining pictures types (neutral: 11.56%, erotic: 10%, babies: 9.84%), as revealed by post-hoc LSD comparisons. Error rates were higher for the short asynchrony conditions (SOA800-late, 12.66, and SOA400, 14.96) than for the long asynchrony one (SOA800-early, 7.66), which indicates again that response errors depended on the picture-target interval.

The results obtained in this experiment replicate the pattern of facilitation and interference observed in Experiment 1 for mutilation images relative to neutral ones. Commission errors were higher for mutilation images than for the rest of the images.

General discussion

In two experiments we investigated the effects of viewing task-irrelevant emotional pictures on the speed and accuracy of a subsequent non-emotional go/no-go task. With a short (200 ms) picture-target asynchrony, mutilation pictures induced a slowdown of reaction times both for key press (Experiment 1) and key release (Experiment 2) responses. Critically, with a longer (600 ms) picture-target asynchrony, facilitation rather than interference was observed for mutilation pictures (Experiments 1 & 2). Moreover, there was a trend for error rates to be higher for mutilation pictures than for neutral ones. Taken together, these results suggest the existence of an immediate (200 ms) emotional effect, interference, as observed in the SOA800-late and SOA400 conditions, and a delayed (600 ms) emotional effect, facilitation, as revealed by the SOA800-early condition.

Mechanisms for the Immediate Emotional Effect

Interference effects have been accounted for by two different mechanisms. First, emotional stimuli, especially those of negative valence, appear to capture attention (Bradley, Mogg, & Lee, 1997; Eastwood, Smilek & Merikle, 2001; White, 1996) thereby reducing the amount of attentional resources allocated to the processing of the target. The slowdown of responses and the error rates observed in the immediate emotional effect (SOA800-late and SOA400, the short asynchrony conditions) are compatible with this hypothesis. It is suggestive that the time interval (200 ms) at which the effect arises matches closely that of attentional blink (Nieuwenhuis, Gilzenrat, Holmes, & Cohen, 2005), that has been shown to occur also when unconditionally aversive stimuli (as were the mutilation pictures in our case) appear briefly before the target, hindering its detection (Most, et al., 2005). Interestingly, this attentional blink effect fades completely for distractor–target SOAs of around 800 ms and longer.

Second, it has also been suggested that interference effects can be the result of the preactivation of motor programs incompatible with the response demanded by the task (Pereira et al, 2004, 2006). According to this idea, positive and negative stimuli should activate approach and withdrawal responses, respectively. There is ample experimental support for this proposal. For example, Duckworth et al. (2002) demonstrated that pulling a lever is faster than pushing it when responding to positive stimuli, but the converse is true when responding to negative ones. In a similar study, Wentura, Rothermund, and Bak (2000) showed that go/no-go lexical decisions tended to be faster when subjects had to press a key in response to a positive word, but also when they had to release the key in response to a negative one. Hare et al. (2005) observed a slowdown of an approach response (keystroke) when subjects responded to fearful faces, and that the activity in the amygdala correlated well with reaction time. In our procedure the preactivation elicited by the emotional pictures adds to that triggered by the warning cue. However, the 'Emotion analysis' we performed should be free of the effect of the triggering cue –at least if we assume that the effects of the cue and the emotional content on response facilitation or interference are additive– since we are comparing among emotional conditions that share the same trial structure. Therefore, in Experiment 1, a purely motor account would predict response facilitation for positive images relative to neutral ones, and interference for negative pictures, as the response demanded was, if anything, an approach one (keystroke). Conversely, for Experiment 2 such explanation would predict interference for positive images, and facilitation for negative images, as the response demanded there have been considered by some (Wentura et al., 2000) a withdrawal one (key release). Since the effects observed were approximately the same in both experiments, our results favor the attentional account

over the motor pre-activation account as an explanation of the immediate emotional effect.

Mechanisms for the Delayed Emotional Effect

The emotional modulation of attention accounts for the response slowdown in the immediate emotional effect (SOA800-late and SOA400 conditions), but it seems difficult to accommodate to explain the facilitation of the response in the delayed effect (SOA800-early condition). Apparently, once the effect on attention has gone by, emotion acts by speeding up the behavioral response, irrespective of its direction (key press or key release). This facilitation may be explained by two mechanisms. First, it is well known that emotional stimuli modulate the activity of visual areas in the cortex, (see Pessoa, 2008, for a recent review) possibly by means of direct or indirect projections from the amygdala to the visual cortex (Freese & Amaral, 2005; Phelps & LeDoux, 2005). This effect of emotion on perception has been shown to occur even when the emotional stimulus is task-irrelevant (Phelps et al., 2006). Bocanegra & Zeelenberg (2009) showed that task-irrelevant fearful faces improve the visual sensitivity to some stimuli (low spatial frequency Gabor gratings) but impair the sensitivity to others (high spatial frequency Gabor gratings). Padmala and Pessoa (2008) observed increased sensitivity in a simple visual detection task to targets that had been paired to mild electrical shocks compared to the same physical targets but devoid of affective value (see Lim & Pessoa, 2008, for a similar result using more complex emotional stimuli). Moreover, enhanced detection of affectively charged targets was associated to increased BOLD activity in areas of primary visual cortex, whereas detection of affectively neutral targets had no relation to BOLD activity in the corresponding early visual areas. The areas of increased BOLD activity for emotional

targets matched retinotopically their spatial location. Interestingly, Lim, Padmala and Pessoa (2008) have shown that this enhanced processing for emotional stimuli depends on the attentional resources available. Under low-load attentional conditions, task-irrelevant fearful faces paired to aversive electrical stimulation elicited stronger activations in the amygdala and the fusiform gyrus than unpaired fearful faces. However, neither activation in these brain areas nor behavioral performance were different for both types of stimuli under high-load attentional conditions, indicating that facilitation depends on the availability of attentional resources. Thus, if we assume that the attentional load of our go/no-go task is low, and that attention can be disengaged from the emotional stimuli in a few hundred milliseconds (Koster, Crombez, Verschuere, Van Damme & Wieserma, 2006), our facilitation effect may well be explained by this perceptual enhancement mechanism, given the evidence that affect-laden stimuli, specially mutilation ones, increase visual brain areas' sensitivity to the target.

Emotional processing could also affect later stages of the response selection process, such as decision-making. The speed-up of responses observed in the delayed emotional effect is reminiscent of a recent theoretical proposal (Simen et al., 2006) that posits that average reward rate provides a global signal that controls the decision threshold during the task: the larger the reward rate, the lower the selected decision threshold. This idea has been supported by experimental results indicating that participants adjust their decision threshold in order to maximize reward rate in decision-making tasks (Simen, Contreras, Buck, Hu, Holmes, & Cohen, 2009). It has been suggested that not only the average rate of recently received rewards but also the opportunity to escape from an aversive situation may lower the decision threshold (Niv, Daw, Joel, & Dayan, 2007). Thus, we can speculate that an unpleasant stimulus might

cause the lowering of the threshold for a decision made around that time, consequently speeding up responses regardless of their direction. This explanation, moreover, is consistent with the observed increase of error rates in Experiment 2.

A direct way to contrast these two hypotheses (perceptual enhancement vs decision threshold lowering) could be to present the emotionally charged stimuli in a sensory modality different from that of the targets (e.g. emotional auditory stimuli in a visual discrimination task or affective visual stimuli in an auditory discrimination task). If the effects are reproduced, that would argue against an explanation in terms of perceptual facilitation.

Conclusions

In summary, the emotional go/no-go procedure we have developed here enabled us to show the emotional modulation of behavior when a very short-lived emotional image preceded the target by several hundred milliseconds. We have also shown how this effect depended on the contents of the emotional image and of its onset asynchrony relative to the target. Our first experiment revealed that, compared to neutral images, mutilation images tended to speed up responses when they were displayed 600 ms before the target onset, but they slowed down the responses when presented 200 ms before the target, while babies facilitated the response for the 200 ms delay, but caused no effect for the longer (600ms) one, and erotic images did not differ from neutral ones. The immediate (200 ms) emotional effect can be attributed to a deficit in the processing of the target as attention is still allocated to the emotional image, while the delayed (600 ms) effect may be caused by either an emotion induced perceptual enhancement or a reduction in the response threshold. An alternative possible explanation of the delayed

effect, namely, that the facilitation observed at the longer picture-target asynchrony was caused by the preactivation of a particular motor program by the mutilation pictures has been ruled out by the results of our second experiment, in which we replaced the demanded response by its antagonist (releasing a key instead of pressing it), but reproduced the results of Experiment 1. The overall results presented here agree with previous findings (Coombes, Cauraugh & Janelle, 2007; Most et al., 2005) and further suggest that emotion can modulate typically "cognitive" processes, such as attention and decision-making (McClure, Botvinick, Yeung, Greene & Cohen, 2007; Pessoa, 2008).

References

Becker, M. W. (2009). Panic search: fear produces efficient visual search for nonthreatening objects. *Psychological Science, 20*(4), 435-437.

Blanchard, D. C. & Blanchard, R. J. (1988). Ethoexperimental approaches to the study of emotions. *Annual Review of Psychology, 39*, 43-68.

Bradley, B. P., Mogg, K., Lee, S. C. (1997) Attentional biases for negative information in induced and naturally occurring dysphoria. *Behavior Research & Therapy, 35*, 911–927

Bocanegra, B. R., & Zeelenberg, R. (2009). Emotion improves and impairs early vision. *Psychological Science, 20*(6), 707-713.

Boulinguez, P., Jaffard, M., Granjon, L., & Benraiss, A. (2008). Warning Signals Induce Automatic EMG Activations and Proactive Volitional Inhibition: Evidence From Analysis of Error Distribution in Simple RT. *Journal of Neurophysiology, 99*, 1572-1578.

Buodo, G., Sarlo, M., & Palomba, D. (2002). Attentional resources measured by reaction times highlight differences within pleasant and unpleasant, high arousing stimuli. *Motivation and Emotion, 26 (2)*,123-138.

Coombes, S. A., Cauraugh, J. H. and Janelle, Ch. M. (2007) Dissociating motivational direction and affective valence: Specific emotions alter central motor processes. *Psychological Science,* 18,11,938-942.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25*, 215-224.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30*, 412-431,

Dolan, R. J. (2002). Emotion, Cognition, and Behavior. *Science, 298*, 1191 - 1194

Duckworth, K. L., Bargh, J. A., Garcia, M., Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological Science, 13(6)*, 513-519.

Eastwood, J. D., Smilek, D., Merikle, P. M. (2001) Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception & Psychophysics, 63,* 1004 –1013

Freese, J. L. & Amaral, D. G. (2005). The organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. *Journal of Comparative Neurology, 486*, 295–317.

Gómez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General, 136*, 389-413.

Goldstein, M., Brendel, G., Tuescher, O., Pan, H., Epstein, J., Beutel, M. et al. (2007) Neural substrates of the interaction of emotional stimulus processing and motor inhibitory control: An emotional linguistic go/no-go fMRI study. *Neuroimage, 36*, 1026–1040.

Hare, T. A., Totteham, N., Davidson, M. C., Glover, G. H., & Casey, B.J. (2005). Contributions of amydala and striatal activity in emotion regulation. *Biological Psychiatry, 57*, 724-632.

Irtel, H. (2006). *PXLab: The Psychological Experiments Laboratory* [online]. Version 2.1.9. Mannheim (Germany): University of Mannheim. Available from <http://www.pxlab.de>.

Koster, E. H. W. , Crombrez, G., Verschuere, B., Van Damme, S. & Wiersema, J. R. (2006). Components of attentional bias to threat in high trait anxiety: Facilitated engagement, impaired disengagement, and attentional avoidance. *Behaviour Research and Therapy, 44,* 1757–1771

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry, 44*, 1248-1263.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-6*. University of Florida, Gainesville, FL.

Lim, S. L. & Pessoa, L. (2008). Affective learning increases sensitivity to graded emotional faces. *Emotion, 8*, 96-103.

Lim, S. L., Padmala, S., & Pessoa, L. (2008). Affective learning modulates spatial competition during low-load attentional conditions. *Neuropsychologia, 46*, 1267-1278.

LeDoux, J. E., & Phelps, E. (2000). Emotional networks in the brain. In Lewis, M. & Haviland-Jones, J. M. (2000). *Handbook of emotions, 2nd edition*. New York, USA: The Guilford Press.

McCall, C., Tipper, C.M., Blascovich, J., & Grafton, S.T. (2011). Attitudes trigger motor behavior through conditioned associations: neural and behavioral evidence. *Social Cognitive and Affective Neuroscience.* Advance online publication. doi:10.1093/scan/nsr057

McClure, S. M., Botvinick, M. M., Yeung, N., Greene, J. D., & Cohen, J. D. (2007). Conflict monitoring in cognition-emotion competition. In Gross, J. J. (Ed.), *Handbook of Emotion Regulation, 204-228*. New York, USA: Guilford Press.

Megías, A., Maldonado, A., Cándido, A., & Catena, A. (2011a). Emotional modulation of urgent and evaluative behaviors in risky driving scenarios. *Accident analysis & prevention, 43(3)*, 813-817.

Megías, A., Maldonado, A., Catena, A., Di Stasi, L. L., Serrano, J., & Cándido, A. (2011b). Modulation of Attention and Urgent Decisions by affect-laden roadside advertisement in Risky Driving Scenarios. *Safety Science, 49*, 1388–1393.

Moltó, J., Montañés, S., Poy, R., Segarra, P., Pastor, M. C., Tormo, M. P., Ramírez, I., Hernández, M. A., Sánchez, M., Fernández, M. C., & Vila, J. (1999). Un nuevo método para el estudio experimental de las emociones: el International Affective Picture System (IAPS). Adaptación española. *Revista de Psicología General y Aplicada, 52*, 55-87.

Most, S. B., Chun, M. M., Widders, D. M., & Zald, D. H. (2005). Attentional rubbernecking: Attentional capture by threatening distractors induces blindness for targets. *Psychonomic Bulletin and Review, 12*, 654–661.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin, 89 (1)*, 133-162.

Nieuwenhuis, S., Gilzenrat, M. S., Holmes, B. D., & Cohen, J.D. (2005). The role of the locus coeruleus in mediating the attentional blink: A neurocomputational theory. *Journal of Experimental Psychology: General, 134*, 291-307.

Niv, Y., Joel, D., & Dayan, P. (2006) A normative perspective on motivation. *Trends in Cognitive Sciences, 10(8)*, 375-381.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology, 191(3)*, 507-520.

Nobre, A. C., Correa, A., & Coull, J. T. (2007). The hazards of time. *Current Opinion in Neurobiology, 17*, 465-470.

Padmala, S. & Pessoa, L. (2008). Affective learning enhances visual detection and responses in primary visual cortex. *The Journal of Neuroscience, 28* (*24*), 6202–6210

Pereira, M. G., Volchan, E., Oliveira, L., Machado-Pinheiro, W., Rodrigues, J. A., Nepomuceno, F. V. P. & Pessoa, L. (2004). Behavioral modulation by mutilation pictures in women. *Brazilian Journal of Medical and Biological Research 37*, 353-362.

Pereira, M. G., Volchan, E., Guerra, G., Oliveira, L., Ramos, R., Machado-Pinheiro, W., & Pessoa, L. (2006). Sustained and transient modulation of performance induced by emotional picture viewing. *Emotion, 6*, 624-634.

Pessoa, L (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience, 9*, 148-158.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron, 48*, 175-187.

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. [Research Support, N.I.H., Extramural]

Schultz, K. P., Fan, J., Magidina, O., Marks, D. J., Hahn, B., & Halperin, J. M. (2007). Does the emotional go/no-go task really measure behavioral inhibition? Convergence with measures on a non-emotional analogue. *Archives of Clinical Neuropsychology, 22*, 151-160.

Simen, P., Cohen, J. D, & Holmes, P. (2006). Rapid decision threshold modulation by reward rate in a neural network. *Neural Networks, 19*, 1013-1026.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., & Cohen, J.D. (2009). Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1865-1897.

Solarz, A. (1960). Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. *Journal of Experimental Psychology, 59*, 239-245.

Theeuwes, J., Godijn, R., & Pratt, J. (2004). A new estimation of the duration of attentional dwell time. *Psychonomic Bulletin & Review, 11*, 60-64

Wentura, D., Rothermund, K., & Bak, P. (2000). Automatic vigilance: The attention-grabbing power of approach and avoidance-related social information. *Journal of Personality and Social Psychology, 78 (6)*, 1024-1037

White, M. (1996). Anger recognition is independent of spatial attention. *New Zealand Journal of Psychology, 25*, 30–35.

Zald, D. H., Mattson, D. L., & Pardo, J. V. (2002). Brain activity in ventromedial prefrontal cortex correlates with individual differences in negative affect. PNAS, 99, 2450-2454.

CAPÍTULO 3

The consequences of unexpected emotional sounds on driving behaviour in risky situations

The consequences of unexpected emotional sounds on driving behaviour in risky situations

Di Stasi, L.L. (1), Contreras, D. (2), Cañas, J.J. (1), Cándido A. (2), Maldonado, A. (2) and Catena, A. (2)

1 Cognitive Ergonomics Group, Department of Experimental Psychology, Faculty of Psychology, University of Granada, Granada, Spain.

2 Department of Experimental Psychology, University of Granada, Granada, Spain.

Published as:

Di Stasi, L. L., Contreras, D., Cañas, J. J., Cándido, A., Maldonado, A., & Catena, A. (2010). The consequences of unexpected emotional sounds on driving behaviour in risky situations. *Safety Science*, *48*(10), 1463–1468. doi:10.1016/j.ssci.2010.07.006

Abstract

Recent development of systems for assisted driving has raised questions about what features of the stimuli perceived by a driver may improve driving behaviour and road safety. The present study aimed to uncover whether emotional auditory stimuli can affect risky behaviour in hazardous situations. Forty-nine volunteers rode a motorcycle in a virtual environment and went through a number of preset risky scenarios, some of which were cued by a sound (a beep, a positive emotional sound or a negative sound). Results showed that hearing the beep reduced the frequency of accidents in the upcoming risky situation, while the emotional cues did not. Likewise, the beep induced the drivers to decrease their speed and focus their gaze on relevant areas of the visual field, while the emotional sounds did not. These results suggest that auditory warning systems for vehicles should avoid using emotion-laden sounds, as their affective content might diminish their utility to increase driving alertness. These findings could provide important information for the development of new advanced driver assistance systems and in general for the specification of future Human-Machine-Interaction design guidelines.

Keywords: motorcycle simulator; ADAS; IADS; eye movements; PERCLOS.

Introduction

The automotive industry has recently introduced a number of active and passive systems that have helped to reduce the number of road accidents and their severity (Sanchez et al., 2006). In contrast to the widespread advancement of safety enhancing technologies for passenger vehicles there has been only limited development of intelligent transport systems (ITS) for motorcycles. Although many motorcycle manufacturers are investing extensive resources in motorcycle safety research, developing new safety passive devices such as the airbag system for motorcycles (Deguchi, 2007), the development of ITS for motorcycles has been given comparatively little priority (Bayly, 2007).

One of the most relevant developments in the field of ITS is that the Advanced Driver Assistance Systems (ADAS). ADAS refers to systems that interact with the driver with the main purpose of supporting the driving task on the tactical and operational levels (AIDE, 2005). In this realm, several automotive industry researchers have studied the effects of the temporal and physical features of warning systems on driver performance and attitudes (e.g. Cacciabue, 2007; see AIDE, 2005). An important issue in ITS research is how to develop in-vehicle information systems that can help drivers without creating cognitive conflict. We use the term “cognitive conflict” to denote situations in which different types of signals are presented in a way that hinders the ability of a driver to perceive, attend to, or react to the potential emergency, either because they compete for attentional or perceptual resources, increase mental workload (Wiese and Lee, 2004), or because they create a conflict between incompatible responses (Botvinick et al., 2001), for example, turning one’s head to determine the cause of a passenger’s scream versus looking straight ahead, leading to a impairment of the driver’s performance. Numerous researchers have investigated this question, as ADAS has been one of the most active areas of ITS studies in the last two decades (e.g. Piao and McDonald, 2008). However to the best of our knowledge, however, no study has addressed the effects of sudden, emotionally-laden, unexpected stimuli on driving/riding behaviour.

Only one study has investigated the effect of emotional induction while driving (Pêcher et al., 2009) and just a few have taken into consideration the effects of real ecological and emotional sounds on driver performance (Di Stasi et al., 2008 and Di Stasi et al., 2010). These studies

focused on the interaction between cognitive and emotional processes and the effects of these on risk behaviour.

In an experiment using naturalistic pictures of road scenes, Di Stasi et al. (2008) studied the effect of visual and auditory warnings on the speed to judge whether or not the scenes depicted a situation of impending danger. The warnings used in that study were a picture of a car (visual icon), an auditory icon (the sound of skidding wheels), speech ("look out!"), and an abstract sound, that had no conventional or broadly understood meaning (the same used in this investigation, see below and appendix I). The speech message facilitated the detection of risk in the scene, while both the auditory icon and the abstract sound did not. The slowest reaction times were observed when auditory and visual warnings were presented simultaneously and the scene was risky. The latter effect was independent of the type of auditory cue. Such results indicate that concurrent warnings may impair the ability to detect risks. In emergencies, it seems therefore preferable to have a single warning signal and to avoid the use of simultaneous ADAS. These results are in agreement with those of other authors (e.g. Wiese and Lee, 2004) and can be accounted for by the central bottleneck model (Pashler, 1998; Levy et al., 2006) which states that states that there is a central limitation for selecting between responses associated with different simultaneous tasks.

Other studies have investigated the influence of emotional stimuli in more ecological and complex settings (fixed-base driving simulators). Pêcher et al. (2009) used music clips with different emotional valence (happy, sad, and neutral) to influence driver's behaviour. "Happy" music clips also distracted drivers more significantly than sad and neutral clips, inducing them to decrease their speed and to impair lateral control.

Similar results were found by Di Stasi, et al. (2010). In this study the authors showed that emotional sounds could interfere with safe driving behaviour in a simulated road environment. Drivers' performance on a secondary risk detection task was influenced by the emotional features of the sound: a negative-valenced sound (scream) and a neutral abstract sound led to shorter reaction times than either a positive-valenced sound (laugh) or the absence of any sound when the complexity of the task was medium and high (while driving in different traffic conditions and performing a secondary task).

In addition, several researchers in different domains have found that the emotional value of a stimulus boosts both the speed of information processing (Öhman et al., 2001) and the likelihood that the information will be processed (Anderson and Phelps, 2001; Vuilleumier and Schwartz, 2001). Brain imaging studies also suggest an interaction between emotional stimuli and perception (Lang et al., 1998, Padmala and Pessoa, 2008).

All these results can be understood according to the "risk as feelings" hypothesis, which states that risk behaviour depends on two systems (rational-analytic and experiential-affective) that jointly determine the individual's risky decisions (Damasio, 1994, Lowenstein et al., 2001). The total amount of effort demanded from the two systems is called mental workload (Hamilton et al., 1979; Sheridan, 1979). From the risk as feelings view, responses triggered by emotional stimuli may compete for the control of mental resources and the ongoing behaviour, increasing the mental workload and interfering with the appropriate response. Our study aimed to determine whether, as predicted by the risk as feelings hypothesis, emotional sounds can interfere with the rider's response to a risky situation in a complex and dynamic simulation.

Methods

Forty-nine volunteers (age range 18–41, average age 20 years - 43 females) took part in this experiment in exchange for course credits. All participants had a car driving license (having had their license for 10 months on average) but they were naive motorcycle riders. All had normal or corrected to normal vision and normal auditory capabilities, according to the results of a standard audiometric procedure. They all signed an informed consent form that described the risks of the study and the treatment of personal data. The experiment was conducted in accordance with the Helsinki Declaration.

The study was carried out on the Honda Riding Trainer simulator (HRT; Honda Motor Co.; see Figure 1). The same experimental setting of Di Stasi et al. (2009) was used. The riding scenarios had been designed for teaching dynamic and complex time-critical driving skills, including situational awareness, risk assessment and hazard perception. HRT data were collected at a rate of 30 Hz. Potentially risky situations were caused by vehicles (cars and motorcycles) that entered the scene from different sides or by unexpected obstacles on the

road. Some hazard scenes were preceded by an auditory stimulus (at 75 dB) that lasted 370 ms. The sounds were a 740 Hz double-beep tone (available for download at http://www.ugr.es/~davcr/beep.wav; adapted from Cabrera et al., 2006) or one of two affectively charged voices: a woman's scream and a baby's laugh, selected from the international affective digitized sounds (IADS, Bradley and Lang, 1999 – numbers 277 and 110, respectively).

Figure 1. The Honda Riding Trainer (left): this static low-cost simulator is used mainly to improve hazard awareness and perception skills. An example of a hazardous traffic situation presented to the participants (right). Up: schematic representation of the dangerous situation. The red dot represents the participant vehicle. Down: the same scene from the participant's perspective. A stopped vehicle opens the door blocking the trajectory of the driver.

Eye gaze direction was measured at 60 Hz by means of FaceLAB v. 4.2. eyetracking equipment (Seeing Machines Limited, Canberra, Australia). FaceLAB determines 3D head position and gaze direction based on video signals from two Sony FCB-EX480A gray scale CCD cameras mounted on the dashboard of the HRT. We monitored gaze behaviour based on the theoretical assumption that eye activity is an indicator of whether the user is processing information about driving (Underwood, 2007) and of the driver's attentional state (Velichkovsky et al., 2002).

The participants were told that the aim of the experiment was to evaluate their behaviour in a realistic road environment. They were asked to abide to road rules to the best of their ability. A recorded voice gave the riders instructions in Spanish.

The study consisted of a training session and an experimental session. During the training session the participant completed two simulated scenarios: the first showed how to operate the HRT and the second how to handle the motorcycle in a complex situation. After the training session the participants ran a complete lap of the experimental circuit without traffic or any risky situation. Next, the participants rode two equivalent street courses, with dual carriageway, mixed traffic flow, intersections and eight potential hazard scenes. In each of the two courses there were eight risk situations, totalling sixteen hazard scenes. Auditory stimuli were presented 500 ms before the onset of the risky situation in half of the hazard scenes: an abstract neutral tone sounded in four randomly chosen scenes within course 1 and a scream or laugh in four randomly chosen scenes in course 2 (although the selection of the scenes was random, the order of the sounds was counterbalanced). The remaining scenes were not preceded by any signal. The contingency between hazard scene and sound was 0, because the probability of the hazard given the sound equals the probability of the hazard given no sound. In order to avoid any carry-over emotional effect, emotional sounds were presented only during the second course (Pereira et al., 2006). The emotional value of the auditory signal was therefore manipulated within subjects at two levels: neutral versus affectively charged sounds, plus a control no-sound condition.

We identified all road accidents that occurred in a hazard scene. We then took a time window of 3.2 seconds after the onset of the hazard scene in which the first accident occurred. The election of the 3.2 s value was made because that was the delay of the earliest first accident relative to the onset of the corresponding hazard scene. This period was then divided into eight 400-ms segments. Speed was averaged within each segment. Given that accidents may happen at points having different speed limits, we computed the difference between the actual speed and the current speed limit at each time segment. Finally, in order to better observe speed changes after the sounds, we computed the difference between the speed at each segment and the speed in the segment immediately preceding the warning onset.

Gaze coordinates (relative to the screen) were recorded from the onset of the hazard scene to 3.2 seconds post-scene. For the analysis of gaze coordinates we divided the HRT screen in 16 x 12 squares of 64 x 64 pixels. The variable submitted to statistical analysis was the proportion

of subjects whose gaze pointed at each square at least once along the 3.2 second interval.

Results

The majority (75%) of participants had at least one accident in the two experimental riding courses. Since it is likely that having an accident affected the riding that ensued, we considered only the first accident of each course for the analysis. To compare the different sound conditions on equal terms, we divided the overall number of first accidents in each condition by the number of scenes of that condition, for there were twice as many scenes in the silence condition as in the other two. Thirty-two first accidents occurred in silence scenes, yielding a first-accident-to-scene ratio of 4 for that condition. The ratios of first accidents per scene were 4.75 for emotional scenes and 2.25 for scenes preceded by an abstract sound. In order to perform a statistical test about the equality of those figures, we computed the same ratios for each individual subject (e.g., if a certain participant had the first accident of the second course in an emotional scene, the ratio for the emotional condition would be 1 divided by the number of emotional scenes; if the participant had no accidents at all, the ratios would be all 0). The single-subject ratios were then submitted to a repeated measures non parametric ANOVA, which showed significant effects of the sound condition, $\chi^2(2) = 9.38$, $p<.01$. Post hoc Friedman comparisons showed that first accidents were less frequent after the neutral sound than after Silence, $\chi^2(1) = 9.00$, $p=.003$, and Emotional sound $\chi^2(1) = 5.56$, $p = .018$. No differences were observed between Silence and Emotional conditions, $p = .53$, or between positive and negative emotional sounds, 47.4% and 52.6% of the accidents occurred after Emotional sound, respectively, $p = .41$.

The second important result relates to the changes in speed after the scene onset (Figure 2). Only the first accident was used for this analysis, independently of the riding course where it happened. The speed, as described in the method section, was measured before the accident, in a temporal window of 3.2 s starting at the beginning of the hazard scene. The speed changes were submitted to a 3 (sound type: no sound, neutral and emotional sound, between subjects) x 8 (time segment, within subjects) repeated measures analysis of variance. Note that in this analysis, the type of sound is a between-subjects variable since only the first accident was considered. There were significant main effects of type of sound, $F (2, 38) = 6.212$, $p = .0046$, p-rep = .973, and time segment, $F (7, 266) = 8.062$, $p < .001$, p-rep = .999. The interaction

between the two factors was also significant, F (14, 266) = 4.737, p < .001, p-rep = .999. Simple effects analysis of the interaction showed speed differences between time segments for the neutral and no sound conditions, F(7, 56) = 6.54, p < .001, p-rep = .999, and F(7, 126) = 2.136, p = .045, p-rep = .886, respectively, but not for the emotional condition (p = .623). Henceforth, since no differences between positive and negative valence sounds were observed (suggesting that neither valence nor arousal in isolation were responsible for the emotional effect, but both together) these two levels were collapsed in one factor “emotional sounds”. Post

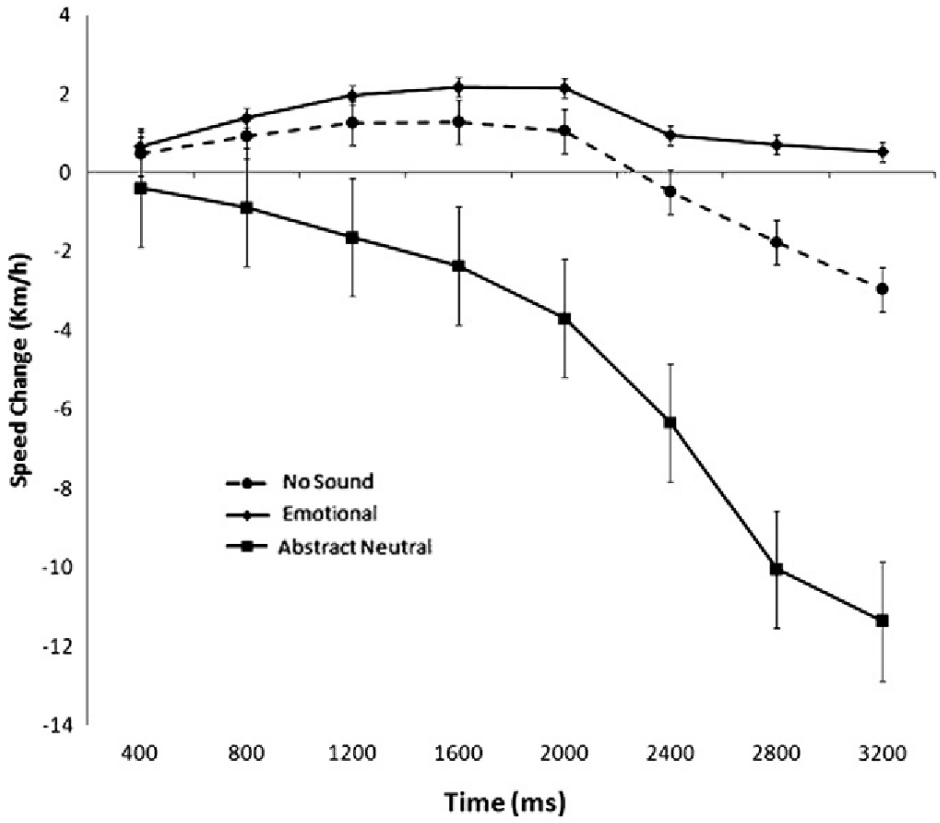

Figure 2. Average speed change after the onset of the risky scene for each type of sound. Standard error of means bars shown.

hoc LSD analysis indicated that participants who had their first accident in a scene that was not preceded by a sound showed a significant reduction of their speed only 2.8 s after the scene onset. However those who had the first accident in a scene that was preceded by the abstract neutral sound reduced their speed earlier, with the decrease reaching statistical significance 2.4 s after the scene onset. Importantly, speed decline was linear for the abstract neutral sound condition (p=0.022), but quadratic for the no sound condition (p = .007). No significant

differences in speed were observed between scenes following emotional sounds and those following no sound but the speeds in these conditions were higher than those in the neutral sound condition in an interval extending from 400 to 2800 ms after the scene onset (all p <.021).

Gaze behaviour is displayed in Figure 3. The proportions of subjects who pointed their gaze to each screen square were submitted to a single proportion z test. Lightness indicates the significance level for each square (see Figure 3 scale). Visual inspection of the statistical parametric maps for each hazard scene revealed that in the emotional and no-sound conditions

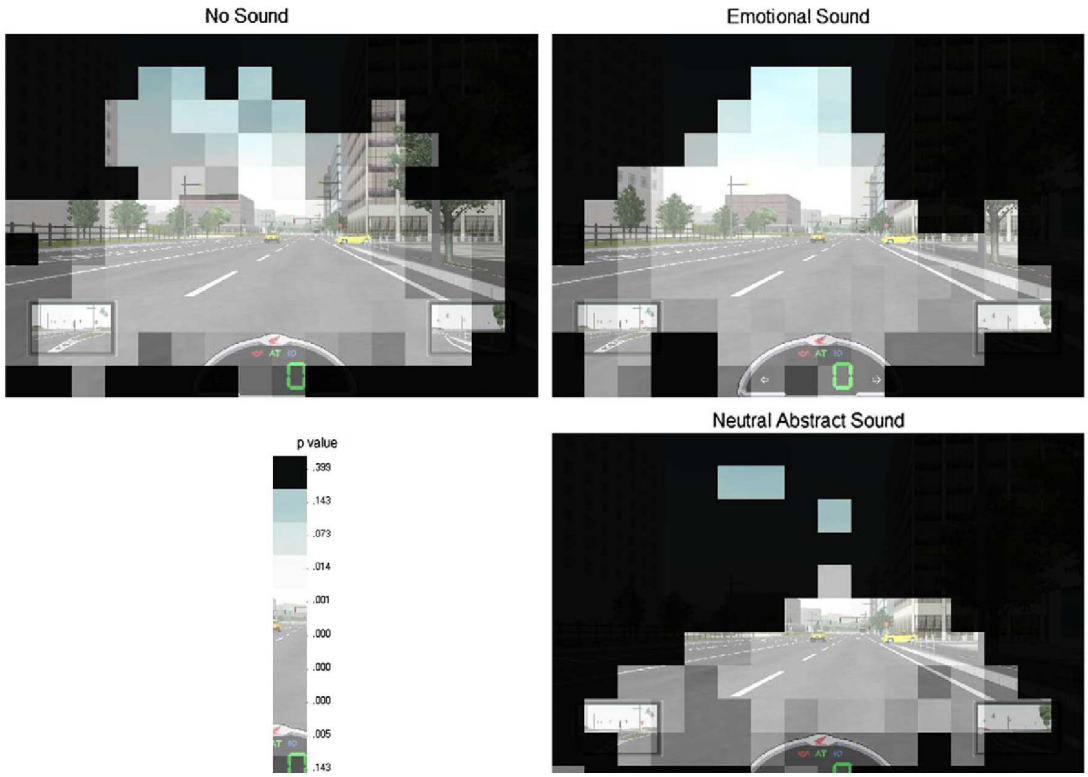

Figure 3. Gaze probability maps during risky scenes for each type of sound.

participants explored both regions relevant to the driving (area of the road in front of the motorcycle, mirrors, etc.) and non-relevant regions (trees, buildings, etc), while after a neutral abstract sound, their gaze tended to focus more on the relevant regions.

Since fatigue might impair performance (Grandjean, 1980) after extended periods of demanding cognitive activity (Lee et al., 2002), we assessed it using the PERCLOS index (Wierwille et al., 1994; Dinges and Grace, 1998), to analyze the effect of time on task. PERCLOS is an index computed upon the duration and frequency of eye blinks, well known indices of time-on-task

and fatigue (Morris and Miller 1996; Schleicher, et al., 2008). The analysis was performed on the first 5000 samples of course 1 (PERCLOS mean: .1432) and the last 5000 of course 2 (PERCLOS mean: .1887) and revealed no differences ($t<1$, $p = .59$). The result suggests that the observed effects are not due to fatigue/vigilance changes between the riding courses.

Discussion

We aimed to test the effects of emotional and neutral sounds on risk behaviour during motorcycle riding in a simulated environment. Two main results were observed. The participants' first accident was less likely to happen in risky scenarios that were preceded by an abstract neutral sound than in those without a sound. However, this decrement in the rate of the first accident did not occur when the signal was an emotional sound (screaming, laughing). On the contrary, emotional sounds tended to increase the likelihood of the first accident. Second, the neutral sound produced an anticipated reduction in speed compared to when the participants entered a risky scenario that was preceded by no sound. Conversely, for hazard situations preceded by emotional sounds, no speed reduction occurred. Given that speed is a fundamental determinant of crash and injury risk (Aarts and van Schagen, 2006), this result is directly relevant to road safety. Moreover, the neutral abstract sound appeared to make the participants focus their gaze on relevant regions of the visual field (the road in front of them, the dashboard of the motorcycle), as compared to the no-sound and the emotional sound conditions.

A possible explanation for these results is that a neutral beep that sounds right before the onset of a hazardous situation may increase the level of alertness, inducing the participants to further allocate processing resources to the situation. This increase in alertness translates into a tendency to direct the gaze towards relevant regions of the scene and into a consequent early reduction in speed, which decreases the likelihood of an accident (Lenné et al., 2008). In risky situations not preceded by a sound, the driver must first evaluate the situation as dangerous for the increase in alertness to occur. Thus, the reduction in speed is delayed and the likelihood of an accident increases. Conversely, when the hazardous situation is preceded by an emotional sound, the activation of the emotional system, as proposed by the risk as feelings hypothesis, can inhibit the recruitment of the cognitive system (Wyble et al., 2007) thereby impairing the

allocation of processing resources to the relevant features of the scene, interfering with the detection of danger, and increasing the risk of an accident. This modulation of attention by emotional stimuli has been widely reported in the literature (Vuilleumier, 2005; Pereira et al., 2006). Alternatively, the emotional stimulus, independently of its affective value and arousal, might trigger an emotional response (e.g.: avoidance/approach) that recruits motor programming resources that are necessary for the response demanded by the current situation.

There are some limitations to our study. First, the lack of rider experience may have exaggerated the effects of the emotional sounds. Among other factors, distraction depends on driver/rider workload (Regan et al., 2008) and, given that workload could be higher for inexperienced drivers, it is plausible that the emotionally-charged sound may have competed for attention, hampering hazard recognition and responding. The neutral tone, instead, probably primed attention (Lee et al., 2002). The effect of the neutral sound (in relation to silence and emotional sounds) might have arisen because the double-beep sounded like what most people would interpret to be a warning, while the emotional sounds did not and were unexpected and out-of-context in the motorcycle simulated environment. This is, however, an interesting result per se, given that our aim was to uncover the effects of unexpected emotional sounds on riding behaviour. More studies are needed to clarify this point.

A second caveat of the present study is the fact that the emotional sounds were always played in the second experimental session; hence the order of presentation is confounded with the emotional value of the sounds. We chose not to counterbalance the order of the sessions to avoid excitation transfer effects. A possible alternative might have been to run both courses in separate days, and counterbalance the order of the courses (emotional-neutral). This could be explored in a future study

A third limitation is related to the emotional features of the selected stimuli. According to the normative ratings for the IADS sounds, the scream (#277) is more arousing (7.86, on a 1-9 scale, Spanish norms; (Fernández-Abascal et al., 2008) than the laughter (#110; 4.71). In our study, the effect on accident rate and speed was similar for the different emotional sounds, but an effect of valence might have been apparent if our subjects had perceived the same arousal

for both sounds. Finally, it is difficult to generalize from a particular positive or negative sound to all positive or negative sounds.

Notwithstanding the above, these results suggest that auditory warning systems for vehicles should avoid using emotion-laden sounds, since adding affective content to the warning sound may reduce its utility as alertness-inducing stimulus. These are conclusions that need further experimental support, however. For instance, in order to confirm that it is the emotional quality of the sounds what limits their utility to increase driver alertness, it would be useful to conduct other experiments in which the neutral sound had more than one level and shared more features with the emotional ones and differed just in its emotional value (i.e. a neutral human voice instead of a machine-produced beep). An alternative study design would include a randomly presentation of all stimuli. Likewise, it would be useful to have concurrent physiological evidence that helped discern whether the effect is mediated by the level of alertness or by some motor effect. It is our hope that this study spurs further research on the effects of realistic emotional stimuli on complex and ecological settings, and encourages the use of basic research in the service of the design of driver assistance systems.

Acknowledgements

This study was supported by a project of the Andalusian Government (Spain) - Excellence Projects PO6-HUM 02375 (A. Catena), P09-SEJ-4752 (A. Candido), by a project of MINCYT PSI2009-12217 (A. Maldonado) and by national grant FPU AP2006-03664 (L.L. Di Stasi). We would like to thank Honda Motor Europe (Aalst office), Honda R&D Europe Italy and Honda affiliate Montesa Honda S.A. of Spain, for their continued support. Thanks also to Lorenzo Angelucci for the technical support and to all members of the Excellence Group for their helpful comments.

Appendix I

Fundamental features of the abstract neutral sound.

Fundamental frequency (Hz)	Harmonics	Signal type	Amplitude (dB)	Pulse rate (Hz)	Pulse length	Silence length	Envelope
740	4	Sine	0	6.3	0.079	0.079	T1

References

AIDE. (2005). AIDE project http://www.aide-eu.org/

Aarts, L., van Schagen, I., 2006. Driving speed and the risk of road crashes: A review. *Accident Analysis and Prevention* 38, 215–224.

Anderson, A.K., Phelps, E.A., 2001. Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature* 411, 305-309.

Bayly, M., Regan, M., Hosking, S. 2007. Intelligent Transport Systems and Motorcycle Safety. Proceeding of the 20th International Technical Conference on the Enhanced Safety of Vehicles. June 18-21, Lyon, France

Bradley M.M., Lang P.J., 1999. International affective digitized sounds (IADS): stimuli, instruction manual and affective ratings. Gainesville, Fl: The Center For Research In Psychophysiology, University Of Florida; 1999 Tech. Rep. No. B-2.

Botvinick, M., Braver, T., Barch, D., Carter, C., Cohen, J., 2001. Conflict monitoring and cognitive control. *Psychological Review 108,* 624-652.

Cabrera, D., Ferguson, S., Laing. G., 2006. Considerations arising from the development of auditory alerts for air traffic control consoles. Proceedings of the 12th International Conference On Auditory Display, London, Uk, 242-245.

Cacciabue, P.C., 2007. *Modelling driver behaviour in automotive environments: critical issues in driver interactions with intelligent transport systems*, Cacciabue P.C. (Ed).Springer-Verlag New York, Inc.

Damasio, *A.R.,* 1994. *Descartes' error: Emotion, reason, and the human brain.* New York: Putnam.

Deguchi, M., Kanbe,S., Hannya, Y. 2007. Basic Research For A New Airbag System For Motorcycle. Proceeding of the 20th International Technical Conference on the Enhanced Safety of Vehicles. June 18-21, Lyon, France

Dinges D.,Grace, R.,1998. *PERCLOS: A Valid Psychophysiological*. *Measure of Alertness As Assessed by Psychomotor Vigilance*. National Highway Traffic Safety Administration TechBrief: FHWA-MCRT-98-006 – 1998.

Di Stasi, L.L., Renner, R., Staehr, P., Helmert, J.R., Velichkovsky, B.M., Cañas, J.J., et al., 2010. Saccadic peak velocity sensitivity to variations in mental workload. *Aviation, Space, and Environmental Medicine 81, 413-417.*

Di Stasi, L.L., Alvarez, V., Cañas, J.J., Maldonado, A., Catena, A., Antolí, A., et al. 2009. Risk Behaviour and mental workload: multimodal assessment techniques applied to motorbike riding simulation. *Transportation Research, Part F* 12, 361-370.

Di Stasi, L.L., Álvarez, V., Serrano, J., García-Retamero, R., Antolí, A., Catena, A., 2008. Visual and auditory ARAS (Advanced Rider Assistance Systems) in road safety decision-making. *International Journal of Psychology 43,* 309.

Fernández-Abascal, E.G., Guerra, P., Martínez F., Dominguez, F.J., Muñoz, M.A., Egea D.A., et al., 2008. The International Affective Digitized Sounds (IADS): Spanish norms. *Psicothema* 20, 104-113.

Grandjean, E., 1980. Fitting the task to the man: an ergonomic approach. 3rd ed. London: Taylor and Francis.

Hamilton, P., Mulder, G., Strasser, H., Ursin, H., 1979. Final report of physiological psychology group. In N. Moray, ed. *Mental workload: Its theory and measurement*. New York: Plenum, 367-377.

Lang, P.J., Bradley, M.M., Fitzsimmons, J.R., Cuthbert, B.N., Scott, J.D., Moulder, B., et al., 1998. Emotional arousal and activation of the visual cortex: An fMRI analysis. *Psychophysiology* 35, 199–210.

Lee, J.D., McGehee, D.V., Brown, T.L., Reyes, M.L., 2002. Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. *Human Factors* 44, 314–334.

Lenné, M.G., Mulvihill, C., Triggs, T., Regan, M., Corben, B., 2008. Detection of Emergency Vehicles: Driver Responses to Advanced Warning in a Driving Simulator. *Human Factors* 50, 135-144.

Levy, J., Pashler, H., Boer, E., 2006. Central interference in driving. Is there any stopping the psychological refractory period? *Psychological Science 17,* 228-235.

Lowenstein G.F., Weber E.U., Hsee C.K., Welch N., 1991. Risk as feelings. *Psychological Bulletin* 127, 267-286.

Morris T.L., Miller J.C., 1996. Electrooculographic and performance indices of fatigue during simulated flight. *Biological Psychology* 42, 343-360.

Öhman, A., Flykt, A., Esteves, F., 2001. Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General* 130, 466-478.

Padmala, S., Pessoa, L., 2008. Affective learning enhances visual detection and responses in primary visual cortex. *Journal of Neuroscience* 28, 6202-6210.

Pashler, H.E., 1998. The Psychology of Attention. Cambridge, Ma: Mit Press.

Pereira, M.G., Volchan, E., Guerra, G., Oliveira, L., Ramos, R., Machado-Pinheiro, W., et al., 2006. Sustained and transient modulation of performance induced by emotional picture viewing. *Emotion* 6, 624-634.

Pêcher, C., Lemercier, C. and Cellier, J.M. (2009). Emotions drive attention: effects on driver's behaviour. *Safety Science*, 47, 1254-1259.

Phelps, E., Ling, S., Carrasco, M., 2006. Emotion facilitates perception and boosts the perceptual benefits of attention. *Psychological Science* 17, 292-299.

Piao, J., Mcdonald, M., 2008. Advanced driver assistance systems from autonomous to cooperative approach. *Transport Reviews* 28, 659-684.

Regan, M.A., Lee, J. D., Young, K. L., 2008. *Driver Distraction: Theory, Effects and Mitigation*. Boca Raton, Florida: CRC Press.

Sánchez, F., Sánchez, D., Paul, A., Baquero, R, Contrera, J., Segovia, M. 2006. Facing Challenges and Requirements for in-vehicle intelligent applications. Proceeding of 13th World Congress & Exhibition on Intelligent Transport Systems and Services. London, 8-12 October 2006.

Schleicher R, Galley N, Briest S, Galley L., 2008. Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? *Ergonomics* 51, 982-1010.

Sheridan, T.B., 1979. Definitions, models and measures of human workload. In N. Moray, ed. *Mental workload: Its theory and measurement*. New York: Plenum, 219–233

Underwood, G., 2007. Visual attention and the transition from novice to advanced driver. *Ergonomics* 50, 1235-1249.

Velichkovsky, B.M., Rothert, A., Kopf, M., Dornhoefer, S.M., Joos, M., 2002. Towards an express diagnostics for level of processing and hazard perception. *Transportation Research Part F* 5, 145-156.

Vuilleumier, P., 2005. How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences 9*, 585–594.

Vuilleumier, P., Schwartz, S., 2001. Emotional facial expressions capture attention. *Neurology* 56*,* 153–158.

Wierwille, W.W., Ellsworth, L.A., Wreggit, S.S., Fairbanks, R.J., Kirn, C.L., 1994. Research on vehicle-based driver status/performance monitoring: development, validation, and refinement of algorithms for detection of driver drowsiness. National Highway Traffic Safety Administration Final Report: DOT HS 808 247, 1994.

Wiese, E.E., Lee, J.D., 2004. Effects of multiple auditory alerts for in-vehicle information systems on driver attitudes and performance. *Ergonomics* 9, 965-986.

Wyble, B., Sharma, D., Bowman, H., 2007. Strategic regulation of cognitive control by emotional salience: a neural network model. *Cognition and Emotion* 22, 1019-1051.

CAPÍTULO 4

Cognitive control ERPs are not involved in the correction of ongoing actions

Cognitive control ERPs are not involved in the correction of ongoing actions

David Contreras[1,2], Alberto Megías[1], Julia Gómez-Cuerva[2], Leandro L. Di Stasi[1], Paloma Mari-Beffa[2], Andrés Catena[1]

1. Mind, Brain and Behavior Research Center. Universidad de Granada. 18011 Granada, Spain.
2. School of Psychology. Bangor University. Gwynedd LL57 2AS, United Kingdom.

Correspondence should be addressed to:
David Contreras
Brain, Mind and Behavior Research Center (CIMCYC)
Campus de Cartuja s/n
Universidad de Granada
18011. Granada. Spain
e-mail: davcr@ugr.es
Tel: +34639704797

Submitted as:
Contreras, D., Megías, A., Gómez-Cuerva, J., Di Stasi, L.L., Mari-Beffa, P. & Catena, A. (submitted). Cognitive control ERPs are not envolved in the correction of ongoing actions. *Nature Neuroscience*.

Abstract

Interrupting the course of ongoing actions and changing the long-term behavioral strategy are separate aspects of the top-down control of behavior. Event-related potential (ERP) studies of cognitive control have identified a negative mediofrontal component peaking ~250 ms after stimulus onset (the cognitive 'N2') that has been linked to cognitive control, but there are conflicting views about its specific role. Here we recorded electromyographic (EMG) and electroencephalographic (EEG) activity concurrently to enable the measurement of the temporal lag between the N2 and the suppression of muscular activity at the single-subject and single-trial levels. The results show that the cancellation of ongoing incorrect responses starts well before the N2 onset. We conclude that the N2, and any posterior cognitive ERP, can only have a role in inter-trial strategy adjustments, not in intra-trial behavior control.

Two capacities of the human brain that have been associated with the concept of "cognitive control" are the ability to interrupt (or inhibit) an ongoing, otherwise automatic action program when an atypical event occurs[1] (e.g., stopping on your tracks if a car rushes out of a garage right in front of you), and the ability to flexibly adapt the future strategy to carry out a task when circumstances change (e.g. paying more attention to garages thereafter). ERP experiments using tasks designed to study cognitive control (e.g., Eriksen flankers task [2 3], stop-signal task [4], and go/no-go task [5]) have pointed to the existence of an EEG component related to cognitive control (labeled "N2" or "N200") which appears when the participant has to overcome a prepotent response tendency, and that manifests itself as a fronto-central negative deflection in the 200–350 ms post-stimulus time-window[6]. However, although the relation of this N2 component to cognitive control is widely accepted, it is unclear whether the N2 is related to the two aspects of cognitive control mentioned above.

It was initially suggested that the N2 is generated by a brain circuit directly responsible for inhibiting prepotent motor responses, since its amplitude was found to be larger under time pressure – which presumably makes it harder to inhibit the dominant response[7]–, when the response requested is more forceful [8], and for participants that commit fewer mistakes [9]. A later, more prominent theory posited that the N2 reflects the emergence of conflict between competing response tendencies (that is, their concurrent activation), with its source located in the anterior cingulate cortex (ACC) [10,11]. Although the original version of this conflict-control loop theory of the ACC postulated that the response conflict in a given trial only affects the level of cognitive control of the trials following it [11], so that the cancellation of an already initiated incorrect response is not influenced by the ongoing detection of conflict, other researchers have explicitly suggested that the level of response conflict might be used by the brain to generate a global inhibitory signal during the current trial that prevents premature responding, giving the decision network more time to gather the information needed to select the correct response [12–14]. In these later views, therefore, conflict drives inhibition. The subthalamic nucleus (STN) is assumed to play the role of monitoring response conflict and producing the global inhibitory signal that blocks premature responses [15,16]. Although no direct relation between STN activity and the N2 is suggested by these theories, the ACC is assumed to monitor conflict simultaneously with the STN in a similar way [14], so the N2 may be regarded as an indirect index of the level of conflict measured by the STN.

The cognitive N2 is hence assumed by some researchers to be an index of neural processes that play a *causal* role in the cancellation of premature incorrect responses. However, causality requires temporal precedence, so this assumption should lead to the prediction that *the N2 will precede the cancellation of suppressed incorrect responses*. And given that the cancellation of premature responses is in some cases an observable phenomenon, manifested by the overt activity of the muscles involved in the suppressed partial response, such prediction is, in principle, a testable one.

No study, though, has checked whether the N2 actually precedes the cancellation of premature responses on a trial-by-trial basis. Burle et al.[17] observed that the N2 in an Eriksen flankers task "invariably starts around the end of the incorrect response activation" (p. 1646) and hence could not be responsible for its interruption, but did not support this naked-eye observation with a statistical analysis. In this study, we recorded EEG and EMG activity simultaneously in order to test empirically whether the N2 precedes the cancellation of partial responses. First, we performed a subject-by-subject ERP analysis in which we assessed the temporal lag between the peak of the single-subject averages of the EMG activity for suppressed responses and the onsets of the N2 components obtained in three different cognitive control paradigms: go/no-go, flankers and stop-signal (Fig. 1). Then, data from the go/no-go task, which produced the clearest N2 component, were submitted to a finer trial-by-trial

analysis to verify that the temporal relation between the N2 and EMG activity, observed in the previous analysis, was also present at the single-trial level.

Results

Behavioral performance and standard ERP analysis

The standard results for the tasks used were replicated, both for behavioral and for ERP measures. Reaction times (RTs), $t(25) = 15.45$, $p < 0.001$, and error rates (ERs), $t(25) = 7.52$, $p < 0.001$, were larger for correct incongruent (461 ms, 15.0%) than for correct congruent (390 ms, 4.90%) trials in the Eriksen task. Error rates increased with the delay between target and stop-signal, $F(4, 100) = 54.35$, $p < 0.001$ (after Greenhouse-Geiser correction). For correct go responses, the average median RT was 305 ms.

Commission error rate was 17.8% on average for the go/no-go task, and 55.3% for the stop task. Side errors occurred in 15.0% of incongruent trials of the flankers task. The average rate of suppressed partial incorrect responses for incongruent trials in the Eriksen task was 34.8%. For no-go trials in the go/no-go task, it was 37.8%. And for stop trials it was 22.2%.

N2 amplitudes were larger for incongruent than for congruent trials of the Eriksen flankers task, $t(25) = 4.53$, $p < 0.001$, and for no-go than for go trials in the go/no-go task, $t(25) = 7.70$, $p < 0.001$. As

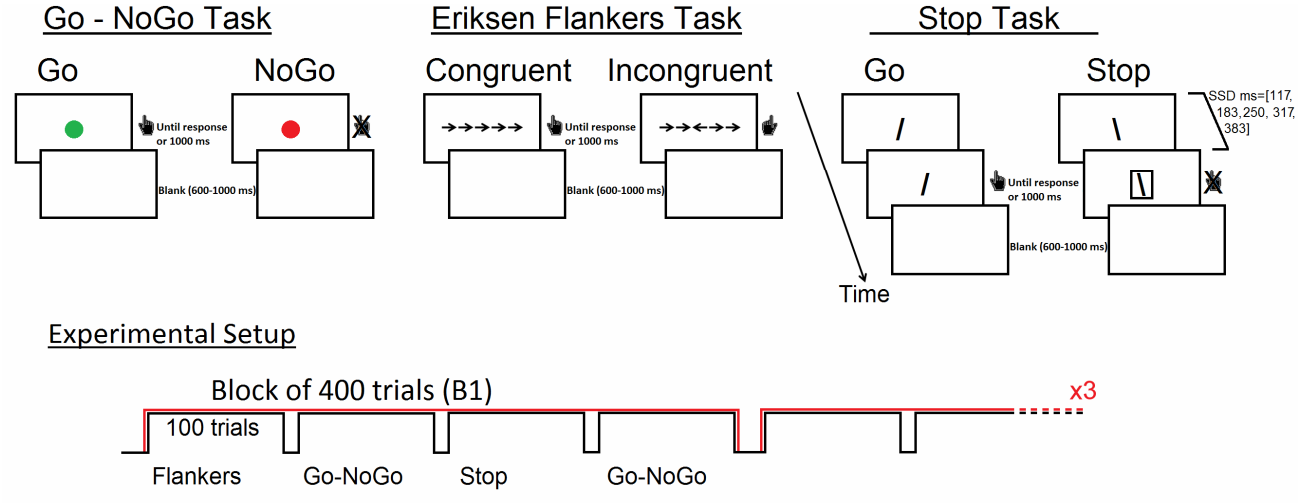

FIGURE 1

Title: schematic description of the experimental tasks used in this study

Legend: Three tasks were used in this experiment. The participant had to carry them out in blocks of 100 trials each, followed by a short resting period, and by a longer resting period every four blocks. The order of the tasks in the first four blocks was counterbalanced across subjects, with the restriction that the go/no-go task had to be in alternate blocks (one possible order is in the figure). The order of the first four blocks was repeated twice, totaling 12 blocks. In the go/no-go task, participants had to press a button to a frequent go stimulus (green circle), and withhold their response to an infrequent no-go stimulus (red circle). In the Eriksen flankers task, the participants had to press the right button when the central arrow pointed right and the left button if it pointed left, ignoring the flankers, which could point in the same direction as the central arrow (congruent trial) or the opposite (incongruent trial). In the stop-signal task, a slash leaning right or left instructed the participant to press the right or left button, but was sometimes followed after a certain delay by a signal (a square outline) telling the participant to withhold her right or left response.

in previous reports[18], larger N2s were obtained for unsuccessful than for successful stop trials, $t(25) = 3.89$, $p < 0.001$.

Subject by subject timing analysis

The N2 onsets determined by segmented regression (Fig. 2)[19] occurred on average after partial responses had already peaked (t_{50}) both at the FCz and Fz electrodes. At FCz, the mean delay was 25 ms in the flankers task, $t(25) = 7.06$, $p < 0.001$ and 11 ms in the go/no-go task, $t(25) = 2.80$, $p = 0.010$. At Fz, the average delay was 18 ms for the flankers task, $t(25) = 3.04$, $p = 0.006$, and 12 ms for the go/no-go task, $t(25) = 2.90$, $p = 0.008$. Note that this means that the cancellation of the response being suppressed had already started by the time of the N2 onset. The areas under the EMG response before the N2 were on average also larger than 50%: flankers task (at FCz, mean = 81%, $t(25) = 7.25$, $p < 0.001$; at Fz, mean = 68%, $t(25) = 2.83$, $p = 0.009$), go/no-go (at FCz, mean = 80%, $t(25) = 3.02$, $p = 0.006$; at Fz, mean = 69%, $t(25) = 3.14$, $p = 0.004$). The delay between t_{50} and the N2 onset as

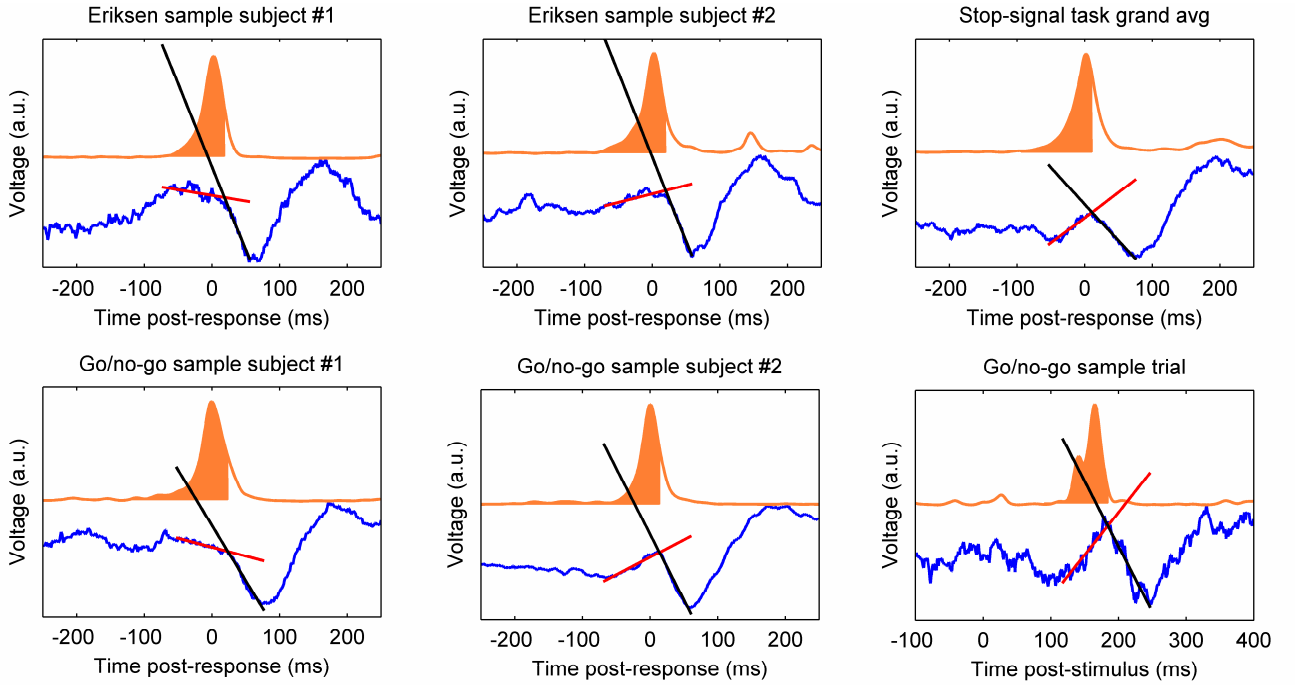

FIGURE 2

Title: Illustration of the method of estimation of the lag between the EMG peak and N2 onset.

Legend: Blue: EEG activity at FCz. Orange: muscular activity. Positive is up. Eriksen graphs were obtained by averaging the activities across all incongruent trials in which the participant made a partial response with the incorrect hand that was successfully suppressed. Go/no-go sample subject graphs were obtained by averaging across no-go trials in which the participant made a partial response correctly inhibited. The Go/no-go sample trial graph depicts the raw EEG and EMG of a single no-go trial with a partial response and a clear N2. The stop-signal graph is a grand average across subjects of trials with partial responses. Activities were time-locked to the partial response time (computed as t_{50}, see Analysis) and then averaged, except for the single-trial graph, that is time-locked to the target stimulus. The x-axis represents time in milliseconds, with time 0 corresponding to t_{50}. The two straight lines were obtained by segmented linear regression in a period of 130 ms preceding the peak of the N2, which is the large negative trough between 0 and 200 ms (or between 200 and 400 ms for the single-trial graph). The onset of the N2 was determined as the cross point of those said lines. The area below the EMG curve and before the N2 onset is shaded in red. EMG and EEG graphs are not to scale. The EMG trace has been shifted up so that it does not overlap with the EEG trace.

measured on the grand-average ERP of the stop-signal task –where a subject by subject analysis was not feasible (see Fig. 3)– was 11 ms, and the area under the EMG curve before the N2 onset was 70% at FCz. Note that all figures and analyses in this section and the next one were done without taking into account the nerve transmission delay. Since the EMG is measured in the forearm, which receives the signal from the brain after a certain delay[20], the EMG traces would be more justly compared to the ERP waveforms by shifting the former backwards in time a few milliseconds, which would increase the size of the effects reported here.

Trial by trial timing analysis in the go/no-go task

The average across subjects of the median lags between t_{50} and the N2 onset was 11 ms, significantly different from zero, $t(25) = 3.14$, $p = 0.004$, and the group average of the area below the EMG and before the N2 onset was 66%, which was significantly different from 50%, $t(25) = 4.52$, $p < 0.001$. Note that these averages need not match the areas and N2 onsets computed in the subject-by-subject analysis reported above, since there the ERP waveform was averaged first for each subject and the onsets and areas computed on the resulting average waveform, while here the N2 onsets and areas were computed on each single-trial waveform first, and then subject medians (for times) and means (for areas) were calculated from the resulting values.

To visualize at a glance the timing of the N2 relative to the muscular activity in partial-response trials, an ERP image of all the trials was generated (Fig. 4), time-locked to the stimulus onset and sorted by t_{50}. The figure highlights the fact that the N2 shifts in parallel with the partial response, and the N2 onset lags the peak of the muscular activity.

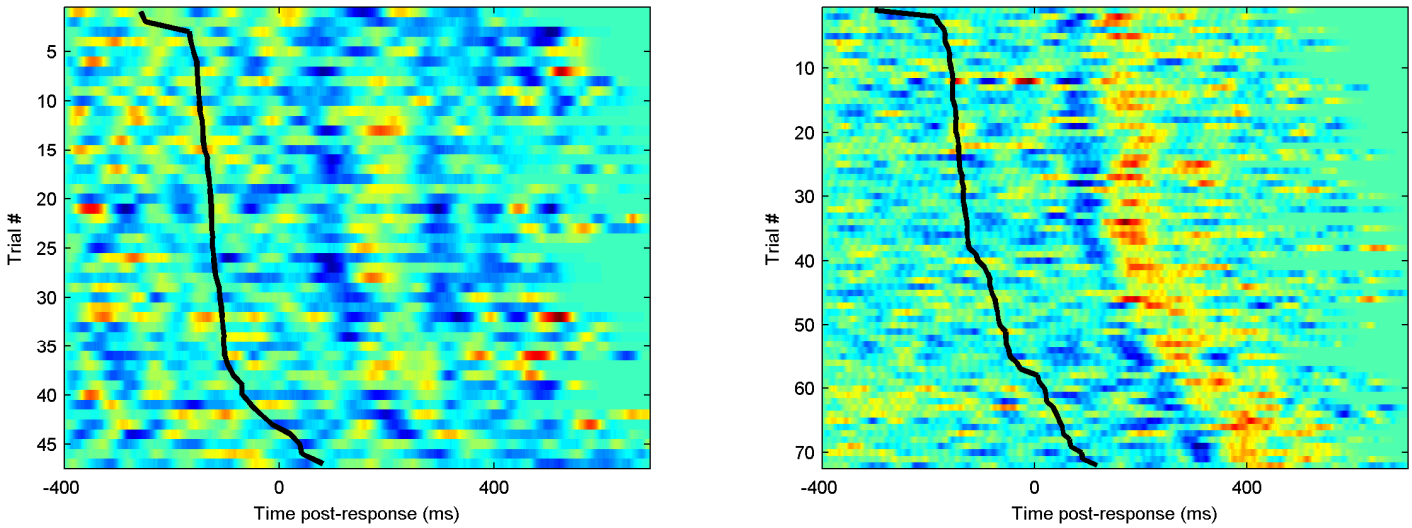

FIGURE 3

Title: Sample ERP images of stop-signal trials with partial response

Legend: Each line represents the ERP of one stop-signal trial in which there was a correctly suppressed partial response. Red: positive voltage. Blue: negative voltage. Each panel shows data belonging to one sample subject that had a relatively clear N2 component. All trials are aligned (at $t = 0$) to the time of the partial response (t_{50}, see Online Methods for details). The black line marks the onset times of the stop signal. The N2 component, visible as a curved blue patch, runs in parallel to the black line, rather than being vertical (as it should be if it were time-locked to the partial response). This makes it unfeasible to estimate the onset of the N2 on the subject average time-locked to the inhibited response, since the N2 will be blurred in that average. A low pass temporal filter was applied for clarity. No across-trials filtering was applied to the image

Discussion

The observation that the onset of the N2 occurs when the cancellation of the motor response has already started is compelling evidence that whatever neural processes the cognitive N2 may index, those processes are not necessary for the cancellation of ongoing incorrect motor responses. This evidence is relevant in light of recent experimental results[16] and theoretical proposals[13,14] suggesting that there is a brain circuit, including neurons in the STN and the inferior frontal gyrus, involved in the active suppression (inhibition) of incorrect responses. Despite the evidence that such mechanism may indeed exist, the current results show that the brain activity signaled by the N2 is, at least, *not necessary* for inhibition, and hence warn against considering the N2 as a direct or indirect electrophysiological correlate of the activity of that inhibitory circuit. The same reasoning can be applied to any cognitive ERP component happening later than the N2, such as the P3.

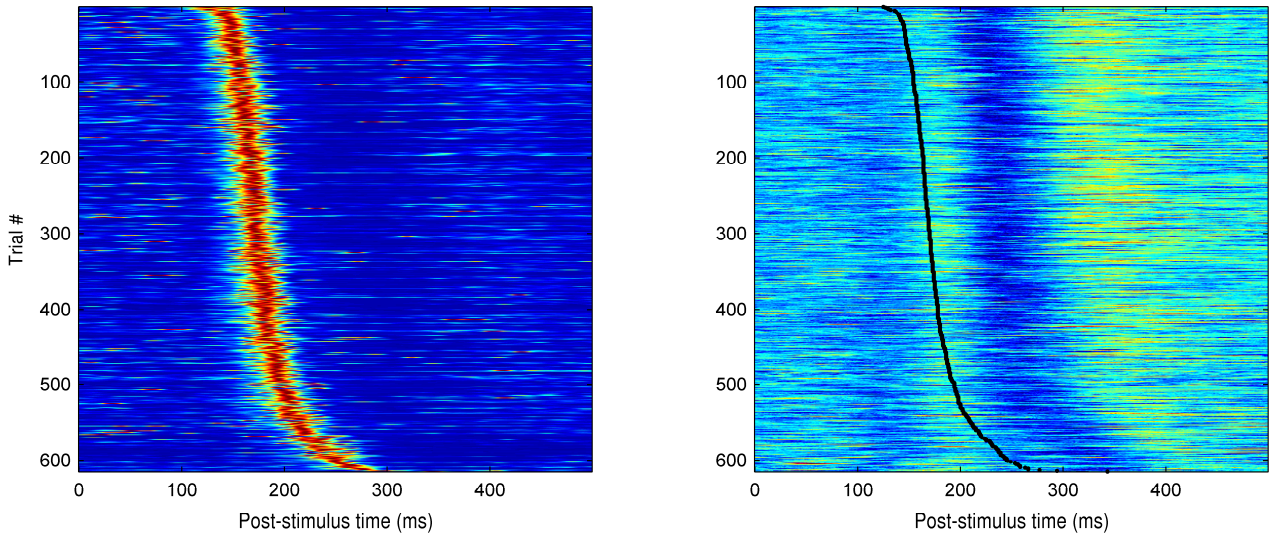

FIGURE 4

Title: ERP image of all correct no-go trials with a clear N2.

Legend: the images include only those partial response trials with a clear N2 component for which an onset could be determined. The left panel shows muscular activity recorded through EMG. The right panel shows EEG activity as measured in FCz. Each trial is represented in one horizontal line. All trials are time-locked to the onset of the no-go stimulus on the screen, and sorted by the latency of the partial response, which was successfully suppressed in every case. On the left panel, red indicates muscular activity, and blue indicates no activity. On the right panel, red means positive voltages and blue, negative voltages. Each trial is normalized so that it spans the whole color range, for visualization purposes. The black dots on the right panel correspond to the time of the partial response represented on the left panel, computed as 50% area latency (t_{50} see Online Methods for details). The N2 components, visible as a blue patch, start after the muscular response for the partial error trials has started to decline. Note that trials from different participants are represented together. No across-trials filtering was applied to the image.

The N2 has also been claimed to reflect the activation of the ACC while it monitors the conflict arising from the co-activation of competing response representations[10,11], a monitoring function that is crucial for the prefrontal cortex to make inter-trial strategic adjustments[21]. Our results are not, in principle, incompatible with this possibility. However, some theorists[14] have gone further in suggesting that the neural signal that informs the ACC about the ongoing conflict between competing responses has a similar counterpart that is fed from motor areas into the STN, and that the STN uses this input to activate the inhibition of motor responses. Simply put, what these latter proposals suggest is that it is conflict monitoring what drives inhibition. If we assume that the emergence of response conflict is

accurately reflected in the N2, what our analyses demonstrate is that the response cancellation starts well before conflict arises. As far as the muscular activity is concerned, the response is, in fact, almost completely cancelled by the time the N2 starts, so the relevance of conflict monitoring for intra-trial motor inhibition, at least from the functional point of view, is questionable.

Although, as said above, our results are not incompatible with the hypothesis that the N2 reflects the monitoring of response conflict, they have some implications either for the connectionist models that support the hypothesis, or for the accepted status of the N2 as an index of ACC activity. The logic of the connectionist models[10] is that, albeit the decision layer is fed by a perception layer in which the stimulus is correctly represented, noise and cross-connections are able to activate the decision unit representing the incorrect response, and sometimes this leads to a wrong response (or a partial one) before the unit representing the correct response is active enough to suppress the activity of the competing wrong response unit by lateral inhibition. In such scenario, it seems quite likely that in most, if not all, partial response trials, the unit representing the correct response, that is more strongly fed by the representation of the stimulus in the perception layer, gets active, at least slightly, right from the beginning of the process, and that its activity grows slowly first, and then faster as it overcomes the lateral inhibition from the competing wrong response unit, that is active mostly by chance (noise and cross-connections). Since conflict is defined in the connectionist models as the co-activation of responses, there should be some conflict from the start of the process, given that both responses are active. That no such conflict is apparent (in the form of some N2 activation) until the incorrect response starts to decrease seems to contradict the model. The simulations ran with the current models support our intuition in this regard (see, e.g., Figure 2 in the study by Burle et al.[17], where simulated conflict starts not just before the peak of muscular activity of partial responses, but even before its onset). Any future proposed model that attributes the N2 to conflict monitoring should be somehow able to predict the observed fact that conflict starts much later than current models predict.

The alternative possibility is to drop the idea that the N2 closely reflects the temporal course of conflict. It might be the case that conflict only gets reflected as a mediofrontal negativity when it exceeds some threshold, so what we observe in the N2 is just the “tip of the iceberg”, and subthreshold conflict cannot be detected in the EEG. That would explain why the onset of the N2 comes later than expected. However, some plausible physiological mechanism should have to be invoked to give account of this “threshold” effect.

The onset of an ERP component is, admittedly, a slippery concept. The estimation of the point at which any component starts is, by necessity, influenced by whatever other components are active at that time[22]. However, some facts seem to support the idea that what we have found may be a robust result. First, given the very procedure we have employed to determine the N2 onset, in most cases we have been taking the peak immediately preceding the N2 (the P2) as the onset of the N2 (see Fig. 2). The common assumption is that most of the amplitude variations at the time of the P2 peak are caused by the P2 itself, not by the N2, which implicitly means that the N2 at that time is essentially non-existent. However, even despite going back in time so much, our onset estimation is still posterior to the peak of the muscular response. Second, if the N2 does in fact extends backwards beyond the peak of the P2, those participants and those trials with larger N2s should have an earlier N2 onset, as a larger N2 would pull the whole curve down and make the peak of the P2 shift backwards in time. However, in our data N2 amplitude did not correlate significantly with N2 onset either across subjects or trial by trial.

At the very least, our results suggest that if the N2 is involved in some way in the cancellation of an incorrect response, its involvement starts (and seems functionally more crucial) long before the N2 peaks. N2 peak amplitude may, hence, have little relation to inhibitory activity.

The simplest alternative interpretation of the N2 is that it reflects some sort of performance monitoring activity in the brain, and there are some prominent theories of ACC function that support this idea[23,24]. These theories share with the original conflict-monitoring hypothesis the assumption that the processes reflected in the N2 affect performance in the trials to come, rather than in the current trial[10,11]. The main contribution of our results is to lend strong support to this notion that the N2 indexes mental processes involved in inter-trial adjustment rather than within-trial control.

Acknowledgments

This work has been partly financed by project PSI2012-39292 of the Spanish Ministry of Economy and Competitiveness and a research scholarship granted to D.C. within the III Plan Andaluz de Investigación of the Andalusian Government.

Author contributions

D.C., A.M., and J.G-C. ran the experiments

D.C., A.M., and A.C. analyzed the data

D.C., A.C., J.G-C., L.D. and P.M-B. wrote the manuscript

Methods

Participants

Twenty-six healthy participants (18 female, 22 right-handed), aged 18–48 (mean 22.2), with normal or corrected-to-normal vision, all of them psychology students at the University of Granada, gave written consent to participate in the experiment, that was approved by the Human Research Ethical Committee at the University of Granada (CEIH).

Experimental procedure

We asked the participants to perform three standard cognitive control tasks: a go/no-go task, an Eriksen flankers task, and a stop signal task. In the go/no-go task, 75% of the stimuli were go stimuli (green circles) and 25% no-go stimuli (red circles). For the Eriksen flankers task, the stimulus was a row of 5 arrows pointing right or left, with the central one acting as target and the other four as flankers. In the stop-signal task, a forward slash required a right-hand response and a backward slash, a left-hand response. The visual stop signal was a squared outline appearing around the slash, lasting 300 ms. The stop signal delay was randomly selected from a set of five possible values (117, 183, 250, 317, 383 ms). All stimuli were displayed on the screen center for 1000 ms or until the participant responded, followed by a blank screen for a random duration between 600 and 1000 ms. Responses were made on a “Serial response box" (Psychology Software Tools, Inc.) using the right and left index finger.

Each participant performed 600 trials of the go/no-go task and 300 of each of the other two tasks, in blocks of 100 trials each, totaling 12 blocks. The first four blocks comprised two of go/no-go trials, and one of each of the other two tasks. The order of the tasks in these first four blocks was counterbalanced across subjects, and the remaining 8 blocks repeated the order of the first four blocks twice. The order was chosen so that the two blocks of the go/no-go task were never consecutive, however.

Electrophysiological recording

We sampled EEG at 1kHz from 61 Ag/AgCl electrodes placed according to an extended 10–20 montage (Easycap - Brain Products GmbH.), using Cz as online reference, and re-referenced to average offline. An EOG electrode was placed under the left eye of each participant, and bipolar EMG electrodes were placed over the abductor pollicis brevis muscles of both hands. Impedances for the EEG and EOG electrodes were below 10kOhm, and below 100kOhm for the EMG electrodes.

Data preprocessing

EEG (including EOG) was high-pass filtered to remove drifts using with a fourth-order (24 dB/octave) phase-shift corrected Butterworth filter with a 0.5Hz cutoff frequency, and we inspected the resulting EEG record visually to manually annotate sections with muscular or drift artifacts. To avoid distorting the EEG record or introducing temporal shifts, we did not apply any low-pass frequency filtering.

EMG was high-pass filtered with 20 Hz cutoff frequency (12 dB/octave). We then rectified the EMG and filtered it low-pass with a 30 Hz cutoff frequency (24 dB/octave).

Classical ERP analysis

We performed a standard ERP analysis on the EEG data to confirm that our results replicated previous findings using these tasks. First, we removed eye artifacts from the EEG record by means of

second-order blind identification (SOBI) ICA[25]. The we obtained single-subject average ERPs time-locked to the target for each condition of interest. ERPs were low-pass filtered (20Hz) and baseline corrected (200 ms pre-stimulus). We measured N2 amplitudes as the average FCz voltage in certain time-windows after target onset (flankers task: 250–350 ms; go/no-go: 200–250 ms; stop task: 200–300 ms). We submitted the amplitudes to paired-samples t-tests with congruence as the factor for the flankers task, stimulus type (go/no-go) as the factor for the go/no-go task and stop success as the factor for the stop-signal task.

N2 onset estimation in single-subject ERPs

The following analysis was applied only to artifact-free trials in which the participant correctly cancelled a partial response (34.8% of trials on average in the flankers task, 37.8% of trials in the go/no-go task, and 22.2% of trials in the stop-signal task).

Partial response detection. We defined "partial response" in one trial as a robust burst of EMG activity ($\geq$ 8 standard deviations (SDs) above baseline) starting no earlier than 20 ms before the imperative stimulus and ending no later than 400 ms post-stimulus onset (or 600 ms for the stop task, where RTs were typically longer). Ten contiguous samples within 2 SDs of the baseline mean before and after the EMG burst defined its onset and offset. Our method was inspired by that of Smid et al.[26].

Partial response cancellation time. We considered the cancellation of a partial response to have started when the amplitude of the EMG burst peaked and began to decrease. For the EEG, it is customary to use the "50% area latency" as a replacement for peak latency, because it is less susceptible to the effects of filtering[22]. Here we similarly computed the area below each partial response EMG curve and above baseline between onset and offset, and then defined the 50% area latency (t_{50}) as the time before which the area under the EMG curve was 50% of the total area.

Estimation of the N2 onset. We then obtained single-subject EEG and EMG activity averages for (a) correct no-go trials time-locked to the partial response, (b) correct incongruent trials of the flankers task time-locked to the wrong hand partial response, and (c) correct stop trials in the visual stop-signal task time-locked to the partial response of the correct hand (i.e., the hand matching the target stimulus). For the go/no-go and flankers tasks, the average ERPs obtained in this way at Fz and FCz had a negative peak around 100 ms after t_{50} whose latency and shape matched those expected for an N2 component. However, for the stop-signal task, there were few participants for which the average ERP showed a clear N2. To further investigate why, and since it has been widely reported that there is an N2 component associated with the stop signal[4], we plotted the potentials at FCz for all partial response trials time-locked to the partial response and sorted by the latency of the stop signal, in an ERP image[27] for each participant (Fig. 3). This revealed that, while in the go/no-go and flankers tasks the onset of the N2 follows that of the partial response (Fig. 4), in the stop-signal task it follows the timing of the stop signal, which explains the absence of a clear N2 in the response-locked averages of the stop-signal task. This made it unfeasible to determine, at the single-subject level, the onsets of the N2 in the stop-signal task using the waveforms time-locked to the partial response.

Thus, we estimated the onset of the N2 for each single-subject ERP waveform of the go/no-go and flankers tasks, but only for the grand-average waveform of the stop-signal task. Following Schwarzenau et. al. [19] we considered the 130 ms trace preceding the N2 peak in each waveform to be composed of two roughly linear segments, with different slopes, and we estimated the onset of the N2 component as the time before the peak when the slope of the curve changed, i.e. when the second segment (the one that ended at the N2 peak) started. In a few cases (5 subjects in the analysis of the flankers task at the Fz channel), we had to shorten or lengthen the 130 ms interval slightly because the two straight segments were too short or two long to fit inside. We located the N2 peak as the minimum

of the ERP trace in the first 200 ms after t_{50}. We took the N2 onset to be the breakpoint in the segmented linear regression.

EMG areas before N2 onset. We also computed the area below the average EMG curve and to the left of the N2 onset for each subject and task, and expressed it as a percentage of the total area below the average EMG curve between the onset and offset of the EMG response. If the N2 onset happened later than the EMG offset, we gave the area the value 100%, and 0% if it occurred before the EMG onset. The parameters to determine the onset and offset of the EMG response in the subject average ERP were slightly different from those used for single trials, since the average is smoother than any single-trial waveform: 15 SDs above baseline defined the response, which we searched for between -200 and 400 ms relative to the partial-response time-lock, and 10 contiguous samples within 3 SDs from baseline determined the onset and offset of the response.

Trial by trial analysis in the go/no-go task

To further check our results, we carried out a trial-by-trial analysis on the go/no-go task. The analysis was restricted to successful no-go trials (i.e., without a key press) where a partial response was observed.

Rejection of trials without N2. Since cognitive control is not necessarily recruited in all trials, we only analyzed those showing a clear N2. We used a measure of similarity between the single-trial and the average N2 waveform to decide whether an N2 was present. Since the single-trial N2 can be delayed or ahead of the average N2, we computer cross-correlations between each single-trial waveform and the average ERP for all time lags between –80 and +80 ms, in a window starting 130 ms before the average N2 peak and ending 50 ms after it. Trials with all r's < 0.71 ($r^2 < 0.5$) were discarded. This was more stringent than simply accepting all trials with their best r statistically different from zero, as doing so would have meant including trials with little similarity to the N2 ERP average.

Estimation of the N2 onset time. We identified the single-trial N2 peak as the minimum of the time-shifted single-trial FCz waveform in a ±30 ms interval around the peak of the average N2 of that subject. We then performed a segmented linear regression[19] on the 130 ms epoch of single-trial potentials preceding the N2 peak, as described in the previous section. We took the breakpoint in the segmented linear regression as the N2 onset.

Estimation of the N2 onset uncertainty. Just as some ERP waveforms did not contain a clear N2, others did not show a sharp change of slope before the N2 peak, so the estimate of the onset time in these cases had a large uncertainty. To obtain a quantitative measure of this uncertainty, 95% we computed confidence intervals for each trial N2 onset by bootstrapping[28], and those trials with a confidence interval larger than 10 ms were discarded. Of the 1475 correct no-go trials with a partial response recorded across all subjects, 615 passed both criteria (similarity and low uncertainty) for inclusion in the analysis.

EMG areas before N2 onset. We also computed the area below the EMG curve before the N2 onset for each trial and expressed it as a fraction of the total area.

References

1. Aron, A. R. *et al.* Converging evidence for a fronto-basal-ganglia network for inhibitory control of action and cognition. *J. Neurosci.* **27,** 11860–11864 (2007).
2. Coles, M. G., Gratton, G., Bashore, T. R., Eriksen, C. W. & Donchin, E. A psychophysiological investigation of the continuous flow model of human information processing. *J. Exp. Psychol. Hum. Percept. Perform.* **11,** 529 (1985).
3. Gehring, W. J., Gratton, G., Coles, M. G. & Donchin, E. Probability effects on stimulus evaluation and response processes. *J. Exp. Psychol. Hum. Percept. Perform.* **18,** 198–216 (1992).
4. Kok, A., Ramautar, J. R., De Ruiter, M. B., Band, G. P. & Ridderinkhof, K. R. ERP components associated with successful and unsuccessful stopping in a stop-signal task. *Psychophysiology* **41,** 9–20 (2004).
5. Pfefferbaum, A., Ford, J. M., Weller, B. J. & Kopell, B. S. ERPs to response production and inhibition. *Electroencephalogr. Clin. Neurophysiol.* **60,** 423–434 (1985).
6. Folstein, J. R. & Van Petten, C. Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* **45,** 152–170 (2008).
7. Jodo, E. & Kayama, Y. Relation of a negative ERP component to response inhibition in a Go/No-go task. *Electroencephalogr. Clin. Neurophysiol.* **82,** 477–482 (1992).
8. Nakata, H. *et al.* Higher anticipated force required a stronger inhibitory process in go/nogo tasks. *Clin. Neurophysiol.* **117,** 1669–1676 (2006).
9. Falkenstein, M., Hoormann, J. & Hohnsbein, J. ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychol. (Amst.)* **101,** 267–291 (1999).
10. Yeung, N., Botvinick, M. M. & Cohen, J. D. The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol. Rev.* **111,** 931–959 (2004).
11. Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. Conflict monitoring and cognitive control. *Psychol. Rev.* **108,** 624 (2001).
12. Bogacz, R. Optimal decision-making theories: linking neurobiology with behaviour. *Trends Cogn. Sci.* **11,** 118–125 (2007).
13. Frank, M. J., Woroch, B. S. & Curran, T. Error-related negativity predicts reinforcement learning and conflict biases. *Neuron* **47,** 495–501 (2005).
14. Frank, M. J. Computational models of motivated action selection in corticostriatal circuits. *Curr. Opin. Neurobiol.* **21,** 381–386 (2011).
15. Aron, A. R. & Poldrack, R. A. Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus. *J. Neurosci.* **26,** 2424–2433 (2006).
16. Aron, A. R. The neural basis of inhibition in cognitive control. *Neuroscientist* **13,** 214–228 (2007).
17. Burle, B., Roger, C., Allain, S., Vidal, F. & Hasbroucq, T. Error negativity does not reflect conflict: a reappraisal of conflict monitoring and anterior cingulate cortex activity. *J. Cogn. Neurosci.* **20,** 1637–1655 (2008).
18. Ramautar, J. R., Kok, A. & Ridderinkhof, K. R. Effects of stop-signal modality on the N2/P3 complex elicited in the stop-signal paradigm. *Biol. Psychol.* **72,** 96–109 (2006).
19. Schwarzenau, P., Falkenstein, M., Hoormann, J. & Hohnsbein, J. A new method for the estimation of the onset of the lateralized readiness potential (LRP). *Behav. Res. Methods Instruments Comput.* **30,** 110–117 (1998).
20. Pascual-Leone, A., Valls-Solé, J., Wassermann, E. M. & Hallett, M. Responses to rapid-rate transcranial magnetic stimulation of the human motor cortex. *Brain* **117,** 847–858 (1994).
21. Forster, S. E., Carter, C. S., Cohen, J. D. & Cho, R. Y. Parametric manipulation of the conflict signal and control-state adaptation. *J. Cogn. Neurosci.* **23,** 923–935 (2011).

22. Luck, S. J. *An introduction to the event-related potential technique*. (The MIT Press, 2005).
23. Alexander, W. H. & Brown, J. W. Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* **14,** 1338–1344 (2011).
24. Holroyd, C. B. & Coles, M. G. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* **109,** 679–708 (2002).
25. Hoffmann, S. & Falkenstein, M. The correction of eye blink artefacts in the EEG: a comparison of two prominent methods. *Plos One* **3,** e3004 (2008).
26. Smid, H. G., Mulder, G. & Mulder, L. J. Selective response activation can begin before stimulus recognition is complete: A psychophysiological and error analysis of continuous flow. *Acta Psychol. (Amst.)* **74,** 169–210 (1990).
27. Jung, T.-P. *et al.* Analysis and visualization of single-trial event-related potentials. *Hum. Brain Mapp.* **14,** 166–185 (2001).
28. Stine, R. An Introduction to Bootstrap Methods Examples and Ideas. *Sociol. Methods Res.* **18,** 243–291 (1989).

CAPÍTULO 5

Moving beyond subtraction methods: how to tailor spatial filters to monitor specific psychological processes in the EEG

Separation of EEG components based on previous knowledge about their temporal course of activity

Moving beyond subtraction methods: how to tailor spatial filters to monitor specific psychological processes in the EEG

David Contreras[1,2], Julia Gómez-Cuerva[2], Paloma Mari-Beffa[2], Andrés Catena[1]

1, Brain, Mind and Behavior Research Center (CIMCYC). Universidad de Granada. Spain
2. Psychology Department. University of Wales Bangor. UK

Corresponding author
David Contreras
Brain, Mind and Behavior Research Center (CIMCYC)
Campus de Cartuja s/n
Universidad de Granada
18011. Granada. Spain
e-mail: davcr@ugr.es
Tel: +34639704797

Submitted as:
Contreras, D., Gómez-Cuerva, J., Mari-Beffa, P., Catena, A. (submitted). Moving beyond subtraction methods: how to tailor spatial filters to monitor specific psychological processes in the EEG. *Neuroimage*.

Abstract

One of the biggest challenges in the analysis of the electroencephalogram (EEG) is to separate the signals of interest originated in the brain from other electrical activity. Spatial filtering may increase signal-to-noise ratio by combining the readings of several electrodes. However, most methods in use for the generation of spatial filters are either based on statistics (data-driven) or require experimental designs where mental processes can be assumed to be either present or absent, in an all-or-nothing fashion. These methods do not take advantage of the knowledge, accumulated for decades of research, about the temporal course of mental processes. Here we develop a method that produces spatial filters able to isolate the brain activities of a set of temporally overlapping, gradually changing mental processes provided that we have a previous hypothesis about their temporal course of activation. The method uses multiple regression and linear constraints to generate filters that pick up signals whose temporal profiles are as close as possible to the hypothesized ones. We test the ability of the method to produce suitable filters to isolate motor and cognitive components by applying it to two sets of real EEG data.

1 Introduction

Electroencephalographic (EEG) analysis ultimately aims to assess mental processes. A variety of methods have been proposed to tease apart the EEG markers of the processes of interest from other electrical activity (noise), ranging from the simple classical averaging ERP technique (Luck, 2005) to linear regression (Catena, Houghton et al., 2009), independent component analysis (Lee, Girolami, Sejnowski, 1999), linear classifiers (Parra, Spence et al., 2005) or matching pursuit algorithms (Durka et al., 2005).

All these methods share the assumption that the observed EEG record is a linear mixture of brain and non-brain electrical activity. According to the strategy used to separate the electrical activity of interest from the linear mixture, methods of analysis can be classified into two broad fuzzy categories, data-driven and model-driven ones (Friston, 1998). Data-driven methods (blind source separation techniques) make no a priori assumptions on the temporal course or scalp distribution of a given mental process, but try to isolate them by making general assumptions about the mathematical properties of *any* plausible brain signal (Stone, 2002). Model-driven methods, conversely, try to fit data to previous specific hypothesis about the temporal course of mental processes -which also guide the experimental design- mainly by using the general linear model.

To the latter category belongs the classical event-related approach (ERP) technique (Luck, 2005). ERP averaging is useful when the mental process of interest has a stable temporal relation with the triggering event in a certain experimental task, since processes unrelated to the event will tend to average to zero. The method is rather unspecific, however, since *all* processes related to the event will influence the average, not just the one under study.

ERP averaging is a form of temporal filtering, since the information about EEG activity in different time points is combined in the process. An alternative possibility to improve signal-to-noise ratio is to combine, at each time point, the EEG activity recorded from several electrodes simultaneously, that is, to filter spatially instead of temporally. This strategy has the advantage of preserving the original temporal resolution, which is unavoidably degraded when averaging across epochs (Luck, 2005). If the brain source of the activity we are interested in is spatially stable, and all other ongoing mental processes have spatially separate sources, the chances are that a spatial filter can be designed that is specifically sensitive to our process of study.

Many spatial filtering techniques have been proposed to analyze EEG data (e.g., Jung et al., 2001, Parra et al., 2005), both data-driven and model-driven. For example, the data-driven independent component analysis (ICA) method produces filters such that the signals they pick up are maximally independent in the statistical sense (Makeig et al., 1996). ICA is appealing due to its theoretical neutrality, but its outcome will only be useful to the researcher if the brain process to be isolated can be assumed to be statistically independent from other concurrent ones, an assumption that, while plausible -for example, for muscle or eye artifacts, (Delorme et al., 2007, see also Onton et al., 2006)-, may fail to hold in some cases (for example, when prefrontal/parietal attentional areas modulate the action of areas involved in visual perception, Friston, 1998).

On the other hand, many of the proposed model-driven methods are binary classifiers (Parra et al., 2005; Poolman et al., 2008; an exception is the Functional Source Separation method, Barbati et al., 2006; Porcaro et al., 2009). Classifiers aim to produce a spatial filter whose output signal discriminates maximally between the EEG time-samples belonging to two experimental conditions, such as left hand vs right hand responses. For a certain mental process to be isolated this way, it is necessary that the only (systematic) difference between the two to-be-discriminated conditions is the presence of that particular process. This is a requirement for the typical ERP experiment too: since all processes time-

locked to the triggering event survive averaging, usually the experimental conditions are carefully designed to match one another in every aspect except for the level of engagement of the process we want to isolate. Hence, designing the experimental conditions for an ERP experiment or for the generation of spatial filters by means of classifiers is a problem essentially akin to that already faced by Donders (1868/1969) in the nineteenth century, i.e., finding ways to make sure that two experimental conditions require exactly the same mental process except for one.

However, some processes are hard to isolate by experimental design. For instance, motor response inhibition (Aron, 2007) can only occur, by definition, once the motor response has been partially activated, so there is no way to create a condition with response inhibition but no response preparation. Conversely, it is hard to design two conditions with different intensities of response inhibition but identical level of response excitation, since stronger inhibition will lead to higher deactivation of the motor preparation. Another example is the somatosensory activity associated to motor responses. This makes it almost impossible to use binary classifiers to generate a spatial filter that is sensitive to motor but not to somatosensory activity, for instance.

A second problem posed by the classifier approach is the implicit assumption that in the time window used to train the classifier the process of interest is either present or absent, with a constant level of activity. This is not a limitation when the target application is to build a brain-computer interface, where the output of the filter is required to be digital, by construction. However, it seems difficult to accommodate when the aim is to isolate processes whose level of activity is expected to vary gradually, such as the force applied to press the button or the degree of process activation requested by an increase in the memory set in a working memory task (for example, Champod and Petrides, 2010; Onton et al., 2005).

In this paper, we propose to replace the binary hypotheses implicitly made in subtraction methods by gradual hypotheses that specify the level of activation of each of the processes of interest at every time point. Hence, for k processes, we suggest assuming that their activities evolve according to k known functions

$$y_i = y_i(t), \text{ with } i = 1 \ldots k.$$

In many cases, we will have reasonable guesses about the shape of the (y_i), derived from theory or from empirical data. For instance, we may assume that motor and somatosensory processes have the temporal profile shown in Fig. 1. In this way, although it is still true that there is somatosensory activity every time a motor action is executed, the slight lag between action and perception and their gradual pace of activation make them distinguishable, and allow us to generate spatial filters specific to each process. Since a binary hypothesis is a particular case of a continuous one, our method includes binary classifiers as a particular case.

The approach we develop here is reminiscent of the method of sparse decomposition using a signal dictionary (Zibulevsky and Pearlmutter, 2001). That method assumes that the source activities may be expressed as a sparse combination of elementary signal functions belonging to a (possibly overcomplete) set (a dictionary), and generates the spatial filters so that the expression of the signals picked up by the filters as a combination of the dictionary elements is as sparse as possible. The dictionaries proposed by Zibulevsky and Pearlmutter (2001) are mathematically guided, so their method is data-driven rather than model-driven. What we propose, instead, is a finite, incomplete and model-driven dictionary, for which we assume that the source activities can be expressed with perfect sparseness (each source activity corresponding to one and only one of the elementary signals in the dictionary).

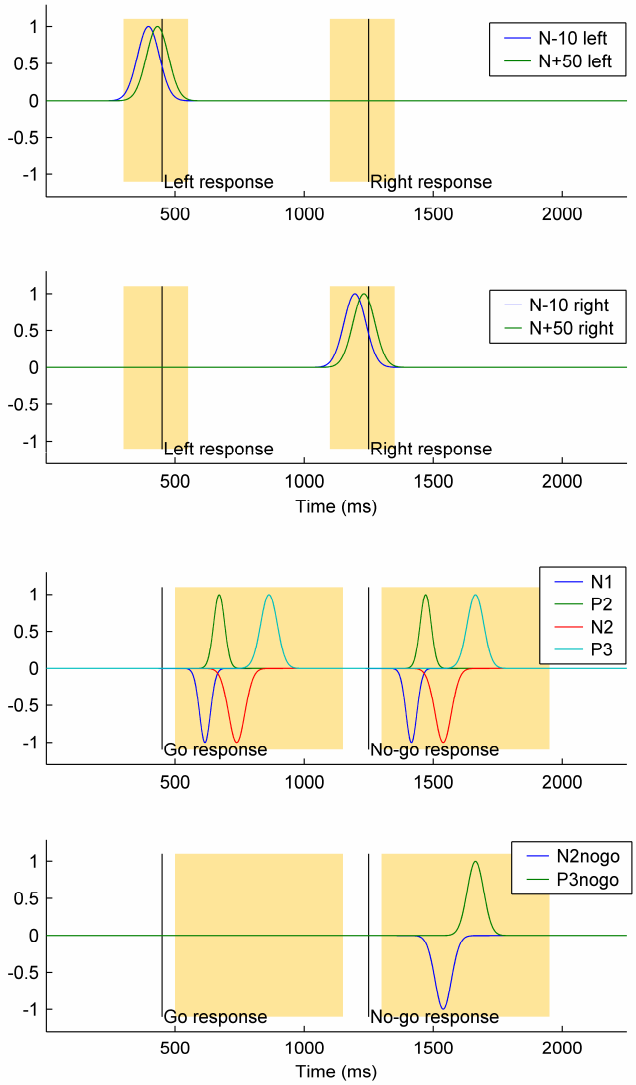

Fig. 1. Graphical depiction of the hypothesized temporal courses of activity used to test the ability of the STOCK method to generate filters selectively sensitive to motor (upper graphs) and cognitive (lower graphs) processes. Left hand processes are assumed to be active only during left-hand responses and vice-versa. Cognitive ("no-go") N2 and P3 components are assumed to be active only in no-go trials of the go/no-go task. The vertical black lines represent stimulus and response triggers in successive trials. The highlighted sections are those used for the multiple regression analysis.

In the next paragraphs we describe the method in detail and then compare its performance with other common spatial filtering methods for disentangling EEG components in two different problems. First we will show the effectiveness of the method for isolating the motor and sensory EEG components related to the action of pressing a button. This is a challenging problem, since sensory and motor components overlap both spatially in the scalp map (because of the proximity of motor and somatosensory cortices) and temporally. Then we will apply the method to the problem of isolating cognitive components in a typical experimental task. In particular, we will describe how to create spatial filters that are selectively sensitive to the cognitive N2 and P3 components.

2 Materials and method

2.1 Problem statement

Given k (possibly overlapping) neural processes of interest, our aim is to generate, for a given EEG record, a set of k spatial filters, each of them sensitive to the activity of a single process and insensitive to the activity of the rest, and all of which are insensitive to noise (muscular, eye blinks, mains) and optimal in the sense that each filter provides the best achievable discrimination between different levels of activity of the process it is meant to detect.

Our problem is, thus, one of creating a method for the Separation of Temporally Overlapping Components based on previous Knowledge about their temporal course (the STOCK method).

2.2 Projection directions and optimal discrimination directions

We assume the scalp potentials at any time point to be a linear mixture of the processes activities $\{y_i\}$ plus noise. Thus, the n-dimensional column vector $\boldsymbol{x}$ containing the n scalp potentials may be expressed as

$$\boldsymbol{x}(t) = \mathrm{A}\,\boldsymbol{y}(t) + \boldsymbol{e}(t) = \Sigma_i\, y_i(t)\, \boldsymbol{a}_i + \boldsymbol{e}(t) \qquad (1)$$

where $\boldsymbol{x}$(t) is the observed EEG waveform, $\boldsymbol{y}(t)$ is the k-dimensional column vector $[y_1(t)\ ....\ y_k(t)]^{\mathrm{T}}$ containing the level of activation of the k processes of interest at time t, $\boldsymbol{e}(t)$ is a noise term and A is the n x k mixing matrix of constant coefficients. The assumption of linearity is made also in ICA and in linear classifiers. (Linearity of the physical medium is implicit in the subtraction analysis of ERPs as well.). ICA assumes both A and $\boldsymbol{y}(t)$ to be unknown, and the $\{y_i\}$ to be maximally independent -in the statistical sense- from each other. Linear classifiers assume $\boldsymbol{y}(t)$ to be known, but binary. By contrast, we assume $\boldsymbol{y}(t)$ to be known and continuous. The implications of assumption (1) are easier to understand if we plot $\boldsymbol{x}$ as a point in an n-dimensional space, in which each Cartesian coordinate represents the potential measured in one electrode. In the absence of noise, the point $\boldsymbol{x}$ representing the readings of the n electrodes is reached by moving y_1 steps along direction $\boldsymbol{a}_1$ and y_2 steps along direction $\boldsymbol{a}_2$. An increase of the activity of the i-th process pushes the point $\boldsymbol{x}$ exclusively along direction $\boldsymbol{a}_i$, which may be thus called the projection direction of the i-th process. $\boldsymbol{a}_i$ is also the scalp map of the i-th process (Parra et al., 2002).

Likewise, a spatial filter $\boldsymbol{w}$ is a set of n weights $[w_1\ ....\ w_n]$ that can be represented as a vector in the n-dimensional space. Applying the filter to a time sample $\boldsymbol{x}(t)$ is done by multiplying the potential $x_j(t)$ at each electrode by the corresponding weight w_j and adding the results up,

$$\hat{y}(t) = \Sigma_j\, w_j\, x_j(t) = \boldsymbol{w} \cdot \boldsymbol{x}(t)$$

that is, by performing the scalar product of $\boldsymbol{w}$ by $\boldsymbol{x}(t)$. Since the scalar product of two orthogonal vectors is zero, a filter $\boldsymbol{w}$ will be completely insensitive to the i-th process if $\boldsymbol{w}$ is orthogonal to the projection direction $\boldsymbol{a}_i$. A filter may be insensitive to several processes simultaneously by being orthogonal to all their projection directions at once.

Since we want to generate a filter for each mental / neural process that is as insensitive as possible to the remaining $k - 1$ processes, we would like to have good estimates of the projection directions $\{\boldsymbol{a}_i\}$ for the k processes. If we rewrite equation (1) as

$$\boldsymbol{x}_i = \boldsymbol{a}_i\mathbf{Y} + \boldsymbol{e}_i \qquad \text{for } i = 1\ ..\ n$$

we see that the expression for the activity of each electrode i is just the general linear model equation. Here the bold face in $\boldsymbol{x}_i$ and $\boldsymbol{e}_i$ indicates they are row vectors containing all time values of voltage and noise for electrode i, respectively. The bold face in $\boldsymbol{a}_i$ denotes a row vector formed with the i-th component of all k scalp maps $\{\boldsymbol{a}_j\}$. Our unknowns are the a_{ij} and the activities $\{y_j(t)\}$ constitute the design matrix (our previous knowledge about the processes of interest). The equation for each electrode is analogous to the equation for the BOLD response of each voxel in standard fMRI analysis. The least-square solution for the a_{ij} is a good estimate for the projection directions $\{\boldsymbol{a}_j\}$ (in fact, the best available estimate if $e_i(t)$ is spherical). Therefore, finding the projection directions amounts just to solving the multiple regression of each $x_i(t)$ on the $\{y_j(t)\}$.

Once we have an estimate of the projection directions $\{\boldsymbol{a}_j\}$, we can move on to generate the set of filters $\{\boldsymbol{w}_j\}$. We would like each $\boldsymbol{w}_j$ to satisfy (a) $\boldsymbol{w}_j \cdot \boldsymbol{a}_m = 0$ for all $m \neq j$ so that the filter for each

process is insensitive to all other processes, and to be such that (b) $\hat{y}_j(t) = \boldsymbol{w}_j \cdot \boldsymbol{x}(t)$ is as close as possible to the hypothetical $y_j(t)$. If we take "as close as possible" to mean "the least squares approximation", then $\boldsymbol{w}_j$ is just the least squares solution to the multiple regression of $y_j(t)$ on the $\{x_i(t)\}$ restricted to the subspace orthogonal to all $\boldsymbol{a}_m$ for $m \neq j$. The direction along which the vector $\boldsymbol{w}_j$ points may be called the optimal "discrimination direction" of the j-th process. In general, the best discrimination direction and the projection direction for a process will not coincide.

2.3 The STOCK method

The solution we suggest to the unmixing problem is, then, as follows:

1. For each of the k psychological processes of interest, formulate a hypothesis about its temporal course of activity: $y_j = y_j(t)$ for $j = 1 \,..\, k$. The function y_j should be defined at least in certain time windows of interest for all j.
2. Estimate the projection directions $\boldsymbol{a}_j$ by computing the multiple regression of each $x_i(t)$ on the set of $y_j(t)$ in those time windows.
3. Project the n-dimensional vector $\boldsymbol{x}(t)$ for each t on the subspace orthogonal to all $\boldsymbol{a}_{m \neq j}$ and then compute the multiple regression of each time series $y_j(t)$ on the set of resulting time series $x_i(t)$. This will produce the optimal discrimination directions $\boldsymbol{w}_j$ restricted to the subspace orthogonal to all $\boldsymbol{a}_{m \neq j}$, i.e. a set of filters $\{\boldsymbol{w}_j\}$ with each filter as sensitive as possible to the corresponding process j and as insensitive as possible to all other processes $m \neq j$.

If we have available good estimates of the projection and discrimination directions of some sources of artifactual noise, such as power line noise and eye blinks (obtained, for example, in some previous BSS analysis of the EEG record), we may first clean the EEG record of those types of noise before performing point 2 (which will give better estimates of the projection directions), and then, in point 3, require the filters to be restricted to the subspace orthogonal to all $\boldsymbol{a}_{m \neq j}$ *and* to the all the known projection directions of artifactual noise. That will produce filters that are also as insensitive as possible to those types of artifactual noise.

2.4 Datasets for testing STOCK

Two datasets were used to test the STOCK method described above. The first dataset comes from an experiment in which 18 participants were asked to perform a stop signal task. On each trial, a forward or backward slash was displayed on the screen center. The participant had to press the right button if the slash shown was a forward slash and the left button otherwise. In half of the trials, a stop signal (a 300 ms, 900 Hz pure tone) was presented to the participant after a short delay. The delay between the target and the stop signal was picked randomly from a set of five equally likely values (83, 167, 250, 317, 384 ms). If the stop signal sounded, the participant had to avoid responding. The target was shown for 1000 ms or until the participant responded. The response was followed by a blank screen lasting a random time interval between 600 and 1000 ms, after which the next target appeared. No feedback was given to the participants.

Each participant performed 600 trials, in blocks of 20 trials. Before each block of trials, a photograph was shown for 4 s. Half the blocks were preceded by a neutral photograph, and the other half by an aversive (mutilation) picture. The blocks preceded by each type of picture were grouped together: either the first 15 blocks were preceded by neutral pictures and the remaining 15 by negative pictures or vice versa (this was counterbalanced across subjects). The original purpose of the experiment was to study the effect of emotional processing on response inhibition, but here we analyze the EEG activity related to sensorimotor processes without any regard to the emotional manipulation. 59-channel EEG was recorded at a 1kHz sampling rate. A vertical EOG electrode was placed under the left eye of each

participant, and bipolar EMG electrodes were placed over flexor digitorum superficialis muscles of both arms, following the location recommendations made by (Zipp, 1982).

The second dataset comes from an experiment in which 17 participants were asked to perform four different tasks: a go/no-go task, an Eriksen task, and two variants of the stop signal task, one identical to the one described above and another with a visual instead of an auditory stop signal. In the go/no-go task the participants were asked to press a button with their right index finger whenever a green circle was shown on the screen center, and to refrain from pressing the button when a red circle appeared. On 75% of the trials the circle shown was green, and on 25% of them it was red. The target remained on-screen until the participant pressed the button or 1000 ms had elapsed. A blank screen ensued, lasting a random time interval between 600 and 1000 ms. Right after the blank screen, the next target was shown. No feedback was given to the participants. In the Eriksen task, a row with five arrows appeared on the screen center on each trial. Each arrow pointed left or right. The participants were asked to focus their attention on the central arrow and press a key with their right index finger if the central arrow pointed right or with their left index finger if the central arrow pointed left. All the arrows flanking the central one pointed in the same direction, which could be the same (compatible trials) or the opposite (incompatible trials) to the direction to which the central arrow pointed. As in the go/no-go task, the target stimulus remained on the screen until the participant pressed the button or 1000 ms had elapsed, and then the screen went blank for 600 to 1000 ms. The two stop signal tasks were identical to the one described above, except for the sensory modality of the stop signal in one of them, which was visual instead of auditory (a squared outline displayed around the stimulus target).

Each participant performed 300 trials of each of the 4 tasks, in blocks of 100 trials, totaling 12 blocks. Each of the first four blocks belonged to a different task. The order of the tasks in these first four blocks was randomized across subjects, and the remaining 8 blocks repeated the order of the first four blocks twice. The order was chosen so that the two types of stop signal task never occurred in consecutive blocks. As to why these four tasks were chosen, the aim was to require the participant to override prepotent or competing responses in different situations (incompatible Eriksen trials, no-go trials of the go/no-go task, and stop trials of the stop signal task) in a single experimental session. 61-channel EEG was recorded, at a 1kHz sampling rate, together with one vertical EOG and two response EMG channels, as in dataset 1.

2.5 Analysis of the datasets

2.5.1 Preprocessing

EEG data were high-pass filtered with 0.5 Hz cut frequency to remove drift, then submitted to the second-order blind identification (SOBI) algorithm (Belouchrani et al., 1997) as implemented in EEGLAB (Delorme and Makeig, 2004). Using custom-made software, power line noise and eye-blink components were identified by visual inspection of the scalp maps and confirmed by visual inspection of the temporal course of their activity. They were then removed from the EEG record. Usually two components (one for eye blinks, one for mains noise) were removed, though sometimes clear higher frequency harmonics of the 50 Hz power line noise were also identified and removed. The scalp maps of the components were later taken as projection directions of the corresponding noise processes. Notice that no low-pass frequency filtering was applied.

Likewise, components of clear muscular activity were identified and removed. A component was considered muscular if its scalp map revealed a projection direction involving just one or two contiguous electrodes, and the output produced by the corresponding filter consisted mostly of long periods of continuous trains of rhythmic spikes. The removal of muscular components was not exhaustive, though, and only the most obvious ones were extracted (5-10 components per participant).

EMG was high-pass filtered (8Hz cut-off frequency) and a band rejection filter centered on 50Hz with a 10Hz bandwidth was applied to it. It was then rectified and low-pass filtered (with 30Hz cut-off).

2.5.2 Dataset 1: Left and right motor components

Several EEG components related to motor activity have been identified in studies measuring scalp potentials and also by subdural recordings (see Shibasaki, 2012, for a recent review). A slowly increasing preparation potential (bereitschaftspotential) starts as early as 2 seconds before movement (Kornhuber and Deecke, 1964). This potential is followed by a steeper one with an asymmetrical scalp distribution, termed the "lateralized readiness potential" (LRP) in cognitive psychology studies (Coles et al., 1988), and starting some 400-500 ms before the peak of muscular activity.

In this paper, however, we focus on two components that appear right before and after the movement. The motor potential (MP or N-10) is localized closely to the site of the movement, contralateral to the limb involved, and peaks 10 ms before the EMG maximum. For its temporal proximity to the movement and its location, it is believed to reflect the activity of piramidal neurons in primary motor cortex (Shibasaki and Hallett, 2006). Its scalp map features a prominent frontal positivity and is negative over parietal electrodes, with the polarity transition slightly anterior to the central line (Gerloff et al., 1998a). A post-movement N+50 component that peaks around 50 ms after the EMG maximum is believed to be related to the somatosensory input caused by the movement (Shibasaki and Hallet, 2006). Its scalp map is frontal-negative and parietal-positive, mirroring the configuration of the N-10, but with a slightly more posterior polarity transition line.

2.5.2.1 Filter and map generation

Using STOCK. Four brain processes were modeled, corresponding to the left and right motor potentials, and left and right post-movement potentials. As a good guess for the the temporal profile of these processes around each response we roughly fitted a gaussian curve to match the width and position of the grand average of the EMG time-locked to the response trigger, and then shifted it 15 ms backwards in time to model the motor component, and 20 ms forward to get the estimated activity for the post-movement component. (Changing the width of the Gaussian to half the original value or assuming a conduction delay of 32+32 ms instead of 15+20 ms had only a negligible effect on the maps and filters obtained.) Fifteen ms is the normal reported motor conduction delay from cortex to muscles in the forearm (Pascual-Leone et al., 1994; Rossini et al., 1994), and 20 ms is the typical latency for the somatosensory cortical evoked potential after median nerve stimulation (Allison et al., 1991). The regressions were performed on 250-ms epochs extending from 150 ms before to 100 ms after each response trigger (Fig. 1).

Using other techniques for comparison. We compared our method to one BSS method (SOBI; Tang et al., 2005), one classifier (logistic regression; Parra et al., 2002), and two prefixed spatial filters that have been specifically proposed for the study of motor potentials, the LRP (Eimer, 1998), and current source density (CSD; Pernier et al., 1988) determined by the surface Laplacian on electrodes C3' and C4', which are located slightly anterior to C3 and C4 (Taniguchi et al., 2001). The classifier was trained to maximally discriminate between right and left hand responses in two 40-ms time windows, one around the estimated peak of the motor component, ("motor" linear classifier), and another around the peak of the post-movement component ("sensory" linear classifier). The CSD was computed by means of two spatial filters generated to yield the best approximation to the Laplacian at C3' and C4' according to the spherical splines interpolation method (Perrin et al., 1989; 1990). C3' and C4' were considered in this analysis to be halfway between C3 and FC3 and between C4 and FC4, respectively.

2.5.2.2 Evaluation of the resulting maps and filters

Dipole fitting for scalp maps. The resulting scalp maps were evaluated for their anatomical plausibility using single equivalent current dipoles fitting. The fit was done for each component of each individual subject using the dipolefitting routine of FieldTrip toolbox (Oostenveld et al., 2011), and the BESA spherical standard head model as specified in EEGLAB v6.01 (Delorme and Makeig, 2004). The exploration of the possible locations was restricted to a grid of 10000 points uniformly distributed in a spherical 4-shell with eccentricities between 0.75 and 0.80 to constrain dipoles to be close to the cortex (giving an interpoint average distance of ~2 mm)

Behavior of the estimated source activities. To evaluate qualitatively the ability of the resulting filters to pick up signals with time-courses similar to those modeled, the filters were applied to the EEG, and the signals they produced were averaged in epochs time-locked to right and left hand key presses separately to create one ERP-like waveform per participant for each component.

Reconstruction of ERPs from source activities. It is interesting to note that in our model the lag between the peaks of the motor and sensory components was chosen to be 35 ms, based on physiological considerations (nerve conduction delay), whereas the reported delay between the peaks of the N-10 and N+50 components in the literature is 60-70 ms (Gerloff et al., 1998a). However, combining two components can easily yield a waveform with peaks at latencies different from those of the peaks of the individual components. To check whether the components we obtained were able to explain the timing of the peaks of the motor potentials documented in the literature, we reconstructed the EEG at Fz using only these components, i.e., we first filtered the EEG using the spatial filters to estimate source activities, and then reconstructed the EEG from these activities by projecting them with the maps. We did three reconstructions for each hand, one using the motor component only, another using only the post-movement component, and a third one using both. Then we calculated, for each hand, the grand average waveforms of the three reconstructed EEG activities at Fz time-locked to the responses made with that hand. To calculate the latencies of the peaks of these waveforms, we first filtered them using a zero-phase low-pass filter with 50Hz cutoff frequency, and then computed the maximum for the motor component waveform, the minimum for the post-movement component waveform (which is frontal-negative) and the maximum and minimum of the waveform generated using both, all of them in a window between -300 and 300 ms relative to the response trigger.

Left-right discriminability analysis. If our filters are selectively sensitive to the motor activity of one particular hand, the signal they produce should discriminate between left and right hand button presses and their ability to do so constitutes a measure of the signal-to-noise ratio in their output. Left-right discrimination performance was measured as the area under the receiver operating characteristic (ROC) curve (Hanley and McNeil, 1982; Parra et al., 2002) computed using the average activities measured by each motor filter in a 40-ms time window around the peak of the modeled motor component. In the comparison, LRP and linear classifiers enjoy a competitive advantage relative to SCD and the filters generated by the STOCK method, given that they just measure the difference in activation between right and left motor areas, while SCD and multiple regression are expected to detect the activity of each motor area separately. For this reason, the activity of the left component was subtracted from that of the right component for SCD and multiple regression before carrying out the comparison. A grand-average ROC curve was generated by arithmetic averaging across subjects of points of equal bias (Lee and Rosner, 2001; Macmillan and Kaplan, 1985)

2.5.3 Dataset 2: Separating cognitive N2 and P3 components

Dataset 2 was used to test the ability of STOCK to generate filters that isolated the EEG markers of suppression/inhibitory processes. The overriding of prepotent responses is commonly related to two

cognitive ERP components: N2 and P3. The cognitive N2 is an EEG component that manifests itself as a fronto-central negative deflection in the 200-350 ms post-stimulus time-window with maximum amplitude around the FCz and Fz electrodes (Folstein & van Petten, 2008). This component appears when the participant has to overcome a prepotent response tendency in order to perform the experimental task correctly. For instance, in the Eriksen flanker task, a larger N2 is visible in incongruent than in congruent trials (Gehring et al., 1992). In the stop-signal task, an N2 is elicited by the stop signal that is not apparent in go trials (Ramautar et al., 2004). In go/no-go tasks, a larger N2 component is observed in no-go trials compared to go trials (Pfefferbaum et al., 1985). There is wide consensus that the cognitive N2 is originated in the anterior cingulate cortex (Karch et al., 2010; Ridderinkhof et al., 2004; Yeung and Nieuwenhuis, 2009).

The P3 is a broad positive ERP component that peaks between 250 and 500 ms post-stimulus, after stimulation in any sensory modality. At least two P3 subcomponents have been identified, one with a parieto-central distribution, called P3b, that has been related to working memory updating, and another related to novelty and the orienting response, called P3a, with a more fronto-central scalp map (Nieuwenhuis et al., 2005). A P3 component associated to cognitive control has also been documented in the literature although it is believed that it may be the same component as the P3a (Polich, 2007). This component usually has a larger amplitude in no-go than in go trials in the go/no-go task (Falkenstein et al., 1999), for incongruent than for congruent trials of the Eriksen task (Clayson and Larson, 2011) and for stop than for go trials of the stop signal task (Kok et al., 2004). Intracraneal recordings have shown P3a-like activity in several cortical and limbic areas (Halgren et al., 1998). Limbic areas comprise the cingulate cortex and the hippocampus, and although lesions of the latter affect P3a amplitude (Knight, 1996), it is unlikely that its activity contributes directly to scalp potentials due to its deep location. In the neocortex, two main clusters of neurons show P3a-like activity: one in the temporoparietal junction (TPJ), and another in the lateral prefrontal cortex (LPFC) (Halgren et al., 1998). Lesions of the TPJ affect both P3b and P3a amplitudes, while LPFC lesions seem to affect the P3a more selectively (Løvstad et al., 2012; Soltani and Knight, 2000).

Although functionally and anatomically dissociable, the amplitudes of the cognitive N2 and P3 tend to covary in many situations, including the experimental tasks we used to generate our dataset 2. Note that the conditions in which the cognitive N2 and the P3 are larger are exactly the same in those tasks, as described above. This relation between the N2 and P3 is so tight in practice that early studies often referred to N2 and P3 jointly as the "N2-P3" complex (e.g., Sams et al., 1983). More recent studies have documented their functional independence (Albert et al., 2013; Enriquez-Geppert et al., 2010; Folstein and Van Petten, 2008), but their tendency to covary renders them *statistically* dependent, which makes it hard for ICA to separate them.

As the dataset used for this analysis comprised data belonging to three different tasks (go/no-go, Eriksen flankers task and Stop-signal task), it provided the opportunity to test whether filters generated using data from one task work appropriately on data from other tasks. To keep things simple in this presentation of our method, we restricted ourselves to modelling the behaviour of the components in the go/no-go task, and then applying the resulting filters to the go/no-go task and also to the Eriksen task as a test of their generalizability. Creating one model for all conditions in all tasks simultaneously would have involved making some too adventurous assumptions about the relative amplitude of the components in the different experimental conditions of the different tasks and about whether their identity was the same or not.

2.5.3.1 Filter and map generation

Using the regression model. The N2 and P3 are typically preceded by other earlier, more perceptual components. Since the best way to make the N2 and P3 filters insensitive to these earlier components was to model them explicitly, we included processes also for the N1 and P2 components.

The N2 and the P3 themselves were divided in two components each: a specific no-go component ("Nogo-N2", "Nogo-P3") that was present only in no-go trials, and a generic one ("N2", "P3") that aimed to capture the other possible N2 and P3 components active during the task but unrelated to cognitive control. Here we borrowed the terminology used for the cognitive N2 and P3 in the context of go/no-go tasks (Gajewski and Falkenstein, 2012). Thus, the model of the psychological activity for this analysis comprised six processes, namely N1, P2, N2, P3, Nogo-N2, and Nogo-P3.

A specific model was generated for each participant, as follows. First, the ERP average waveform for no-go trials was computed for each participant, since it was easier to determine the timing of peaks in no-go trials as they exhibited much larger N2 and P3 components. Then, the typical latency relative to no-go stimulus onset for each of the component peaks was determined using this average waveform, by searching for maxima and minima in certain intervals, as described in Table 1. The correctness of the peaks determined in this way for each participant was checked by visual inspection. Then, for each of the six processes in the model, a flat (zero) level of activity was assumed at all times except when the corresponding component peaked in go/no-go trials. N1, P2, N2 and P3 were assumed to peak in all trials at the asynchronies calculated above relative to the stimulus onset. Nogo-N2 and Nogo-P3, on the other hand, were assumed to peak at those same lags but only in no-go trials. The non-zero activity for each component was modeled as a gaussian profile centered at the average peak latency for that component, relative to the target onset of each go/no-go trial (Fig. 1). The widths of the gaussians used were the same for all participants, but differed between components (Table 1). For the earlier components (N1, P2) the widths were smaller than those of the later ones (N2, P3) in order to mimick the shape of the EEG components in the grand average ERP. Note, however, that the results of the analysis were not very sensitive to the actual widths and shapes used to model the activities of the processes. The amplitude of the gaussians was the same in all go/no-go trials. The regressions were performed only in go/no-go trials, in the sections of the EEG extending from 50 to 700 ms after the onset of the no-go targets.

Component	Peak amplitude, location, and time windows			Gaussian FWHM
N1	min	POz	100-180 ms	50 ms
P2	max	FCz	150-250 ms	50 ms
N2 & N2nogo	min	FCz	200-400 ms	75 ms
P3 & P3nogo	max	FCz	280-500 ms	75 ms

Table 1. Timing and widths of the Gaussian curves used to generate the model for the extraction of cognitive components depicted in Fig. 1.

Removal of motor components. Beside the noise components identified in the preprocessing stage (mains, eye blinks and muscular), in this analysis we were interested in removing the EEG activity related to motor action and post-movement processes before generating the filters for the cognitive components. One possible way to remove them is to model them explicitly. However, since we only

included no-go trials in the regressions, there were few trials with any motor activity (only those with commission errors), and only the activity with the right hand would have been different from zero in the model.

Another way to remove motor and somatosensory processes is to use our knowledge about their projection directions to restrict the filters for the processes of interest to be insensitive (orthogonal) to them by construction. This is what we did in this case. First, we perfomed an analysis identical to that applied to dataset 1 to obtain filters and projection directions for the motor and somatosensory components. The projection directions obtained in this way were then treated in the regression analysis in the same way as those for noise components.

The clean EEG record and the model for the psychological processes were then submitted to the multiple regression analysis to generate one filter for each of the four processes of interest.

Using other techniques for comparison. We compared our method to one classifier (logistic regression; Parra et al., 2002) and one BSS method (SOBI; Tang et al., 2005). To try to isolate the Nogo-N2 and -P3 components using logistic regression, we trained the classifier to discriminate optimally between correct go and no-go trials in two 60-ms time windows, centered around the peak of the N2 and the peak of the P3 in no-go trials, respectively (Parra et al., 2005). The peak latencies were calculated using the single-subject no-go ERPs, as described above. For the BSS method, in order to select the SOBI spatial filters that picked signals most similar to the N2 component, we first computed the temporal average of the signal produced by each of the filters in a 60-ms time window centered around the N2 peak, for all go/no-go trials, and then did the same with the voltages read at FCz. The spatial filter selected was the one whose average activities around the N2 peak had the largest correlation with those in the FCz electrode across go/no-go trials. For P3, the same procedure was followed, but computing the activities in 60-ms intervals around the P3 peak of each participant.

2.5.3.2 Evaluation of the resulting maps and filters

Subtraction of the resulting components from the ERP. Having filters that isolate individual components allows to substract them individually from the EEG activity, by filtering spatially the EEG first to estimate the source activity and then removing this activity by projecting it back to the scalp with the corresponding map. To test the sensitivity and selectivity of the N2 and P3 filters, we compared the average amplitude at FCz in go vs no-go trials in two time windows, one around the N2 peak (260 - 320 ms post-stimulus onset) and another around the P3 peak (350 - 450 ms post-stimulus), and this statistical test was done in three situations: a) with the original EEG record; b) after substracting the Nogo-N2 component; c) after substracting the Nogo-P3 component.

If the filters are both selective and sensitive, removing the Nogo-P3 should not eliminate the difference between go and no-go conditions in the time-window of the N2 peak, but removing the Nogo-N2 should render the difference between go and no-go trials non-significant. Conversely, the difference in the time window of the P3 peak should survive removal of the Nogo-N2 component, but vanish when the Nogo-P3 component was removed. Note that to these predictions underlies the assumption that, apart from the no-go N2 and P3 components, no other components present in the EEG showed a significantly different activity in go vs no-go trials in the time windows chosen for the analysis.

Behaviour of the estimated source activites. Just as we may substract individual components from the EEG, we may also reconstruct the EEG using only the components we want. The reconstructed grand average ERPs containing each of the components in isolation were calculated and plotted for go and no-go conditions. A t-test was performed to compare the amplitude at a 50-ms time window around the peak in go vs no-go trials for each of the components. We expected to observe differences for the

Nogo-N2 and Nogo-P3 components, but no difference for the 'generic' N2 and P3 components or for the more perceptual N1 and P2 components.

Reconstruction of ERPs from source activities. The multiple regression model used to generate the filters with the STOCK method need not be exhaustive, that is, there may exist some background activity, even if it *is* related to the triggering event (stimulus onset in our case), that is not modeled at all, and that doesn't preclude the method from producing good filters. However, it is interesting to see

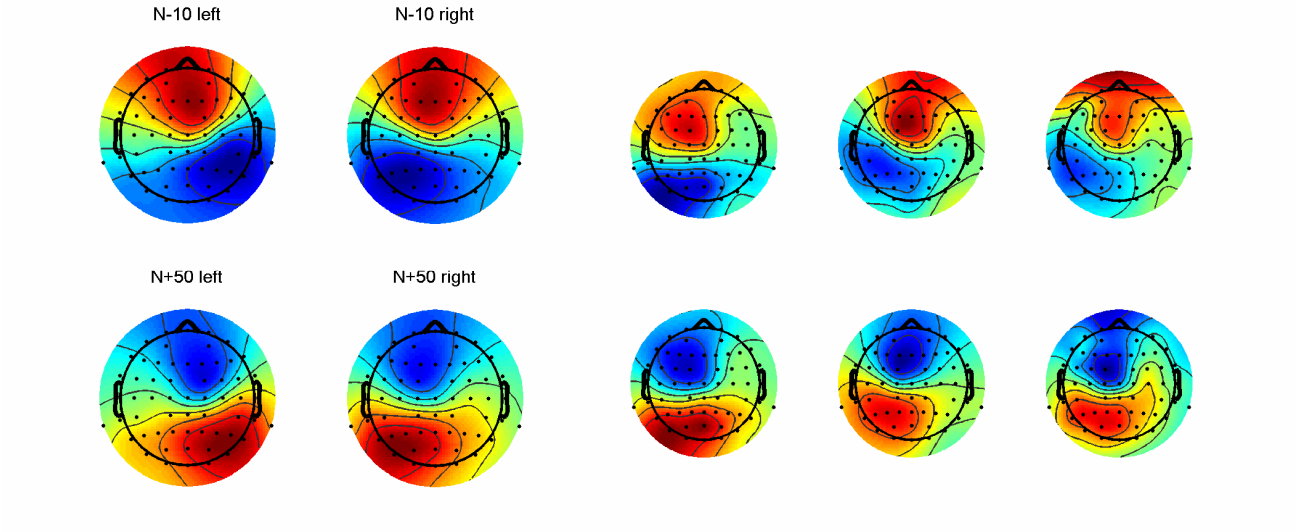

Fig. 2. Scalp maps for left and right motor (N-10) and post-movement (N+50) components, as obtained by the STOCK method. Left hand side (larger maps): grand average scalp maps. Right hand side (smaller maps): scalp maps in three sample subjects of the N-10 right (upper row) and N+50 right (lower row) components. All maps exhibit a clear dipolar distribution suggesting a source located in the central sulcus, contralateral to the response hand. The polarity of motor components suggests a source in the anterior bank of the central sulcus, while that of the post-movement components is consistent with a source in the posterior bank of the central sulcus. The maps were re-referenced to average offline.

whether the model proposed explains most of the event-related EEG activity by comparing the original ERP to the ERP obtained synthetically by combining the signals picked up by all the filters in the set. Thus we projected back all the estimated source activities to the electrode space using the corresponding maps and compared the resulting ERP to the original one.

Application of go/no-go filters to other tasks. There are reasons to expect that the filters generated by our technique using data belonging to one task will not work fine with sections of the EEG from other tasks. Different tasks may requiere different mental processes, and the ability of our algorithm to create spatial filters that are selectively sensitive to the process of interest, even if it is shared across tasks, depends on the presence in the training set of the activity to which the filters should be *insensitive*, so that the algorithm "learns to ignore it". In other words, the filters created by the algorithm tend to be insensitive to the activity unrelated to the processes of interest *that is present in the dataset* used to create them, so there's no guarantee that they will be able to filter out the activity related to mental processes that are only engaged by other tasks.

Anyway, it's an interesting question to what extent the filters obtained with data from one task can be applied in practice to sections of the EEG belonging to different tasks. Since our dataset contained sections recorded in a single session where the participants perfomed different tasks, checking this was possible in our case. We therefore calculated average ERP waveforms for congruent an incongruent Eriksen trials, in which we had substracted the N2 or the P3 components.

3 Results and discussion

3.1 Dataset 1: Left and right motor components

3.1.1 Scalp maps

Grand-average scalp maps and some individual scalp maps generated by STOCK are shown in Fig. 2. The three examples are representative of the variability we found across participants. One remarkable regularity observed across all subjects was the fact that scalp maps showed frontal positive polarity and posterior negative polarity for the motor components, while the opposite pattern was observed for somatosensory components. The polarity transition was located roughly on the central line of electrodes in all cases, slightly posterior for somatosensory maps than for the corresponding motor maps. This was a strong indication that the components we measured had their anatomical origin in opposite banks of the central sulcus, the motor component in the motor cortex (anterior bank), and the post-movement component in the somatosensory cortex (posterior bank). Similar results have been reported elsewhere (Shibasaki, 2012) for maps computed by conventional averaging, both for EEG (Gerloff et al., 1998a) and MEG (Gerloff et al., 1998b), but those maps seemed to be much less clear than the ones shown here. For instance, the scalp map for the motor potential presented by Gerloff et al. (1998a) was a "selective group average" of 10 out 24 possible datasets, made to "enhance the signal-to-noise ratio for the motor potential". By contrast, *individual* motor scalp maps obtained by the STOCK method showed the correct polarity and lateralization for all 18 participants, and almost all of them were clear enough to recognize the general shape outlined above without the need of group averaging.

3.1.1.1 Dipole fitting

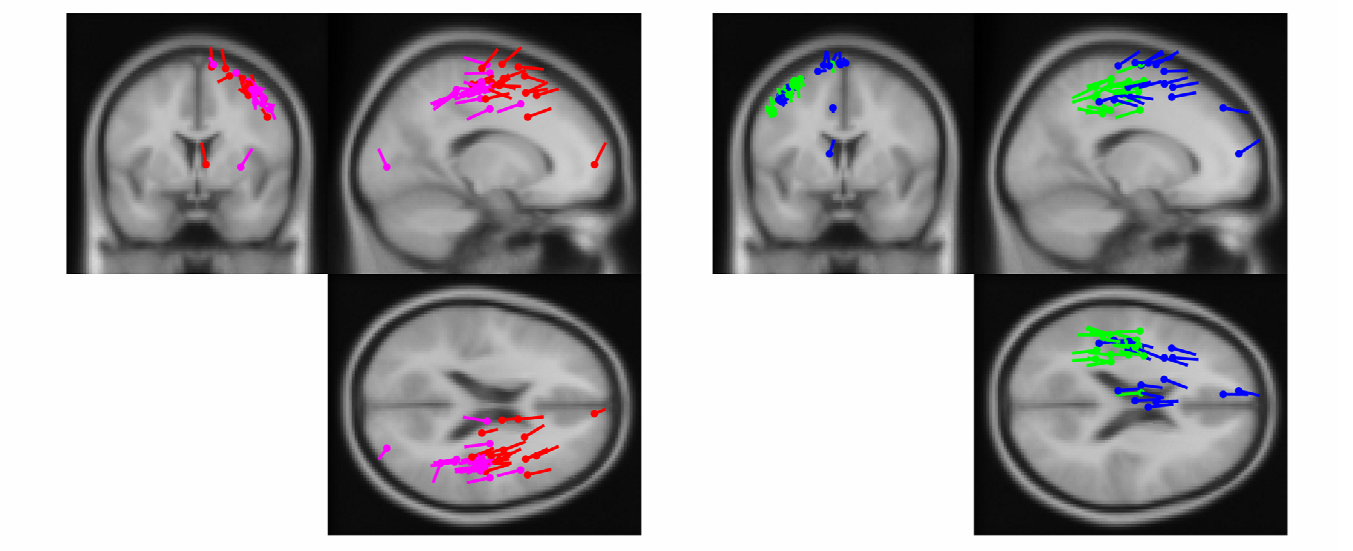

Fig. 3: Locations of the dipoles fitted to the motor and post-movement scalp maps for each of the 19 subjects. Blue = right motor. Green = right post-movement. Red = left motor. Magenta = left post-movement. The cloud of dipoles is centered around motor areas for motor components and slightly posterior (around somatosensory areas) for post-movement components. The dipoles point anteriorly for the motor components and posteriorly for post-movement components, as expected for surface-negative dipoles located in the anterior and posterior banks of the central sulcus, respectively. The dipoles are always contralateral to the response hand.

The locations and moments of the dipoles fitted to the scalp maps are displayed in Fig. 3. The average residual variance for the 72 fitted dipoles (4 for each participant) was 9.87%. The average distance from the center of mass for each type of dipoles was 30 mm for right motor dipoles, 23 mm for left motor dipoles, 14 mm for right sensory dipoles and 19 mm for left sensory ones. The low dispersion of the dipole locations for the post-movement components is noteworthy, although that of the motor dipoles is not so impressive. However, all dipoles were placed on the correct side and had the correct orientation, the motor dipoles pointing anteriorly and the sensory ones posteriorly, which is a considerable achievement for individually fitted dipoles. The position of each motor dipole was anterior to that of the corresponding sensory one in all cases for the left hand and in 16 out 18 participants for the right hand. The cloud of somatosensory dipoles is centered around the hand areas in the central sulcus. The motor dipoles are generally close to motor and premotor areas. Considering the low resolution of dipole localization methods and the anatomical variability between subjects, the fact that the dipoles for each component are near the areas known to control the corresponding process may be considered as sound evidence that the method described in this paper was indeed able to isolate the

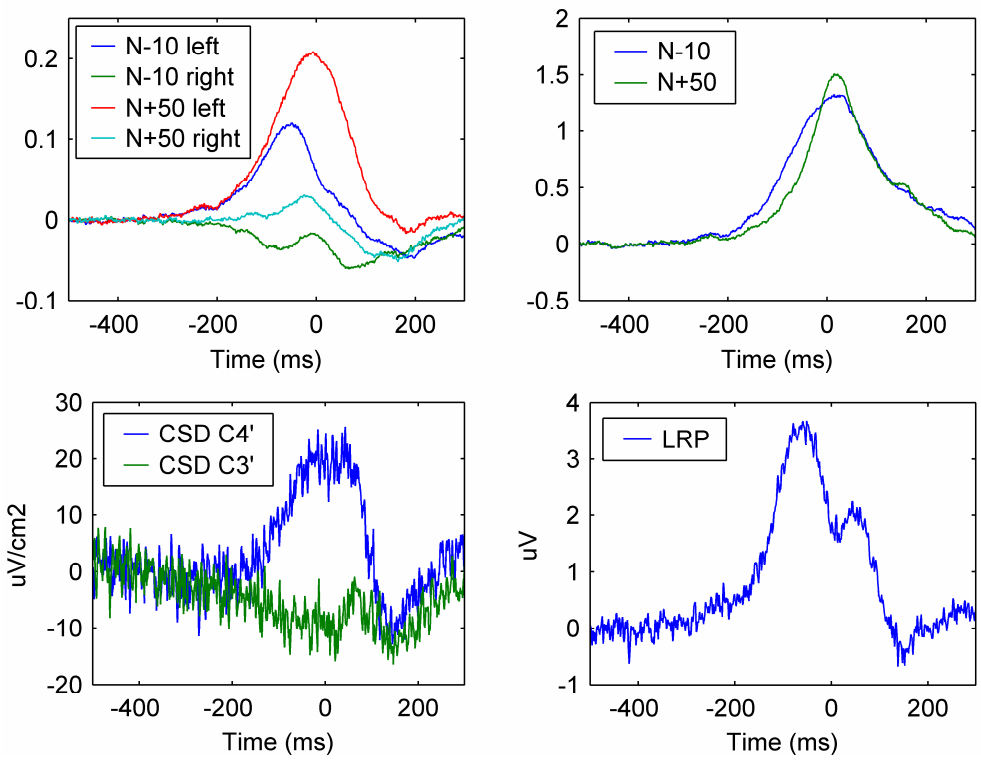

Fig. 4 Grand averages of the activities detected by several filters during left-hand responses, time-locked to the button press trigger. The time of the peak of EMG activity is slightly anterior to that trigger. Upper left: activities picked by STOCK filters. The left motor (N-10 left) component peaks right before the muscular response, and the post-movement (N+50 left) component, slightly later. Right hand filters (N-10 right, N+50 right) detect little activity. Upper right: activities for filters obtained by using logistic regression are unable to separate motor from post-movement components (both waveforms peak at the same time). Lower left: activities measured by current source density (laplacian) filters. The signal is noisy because these are prefixed filters, so they don't adapt to the subject anatomy or the sources of noise. The activity picked by C4' (right hemisphere, left hand) has the temporal course of a somatosensory, rather than motor, component. The C3' filter (left hemisphere, right hand) detects little activity. Lower right: lateralized readiness potential. This is also a prefixed filter, so its output is also noisy. The activity profile corresponds to a motor rather than a post-movement component, but does not separate right from left hand.

desired components.

3.1.2 Filters

3.1.2.1 Temporal profile

Average activity is shown in Fig. 4 for the left motor and sensory components. The overall shape of the component activity resembles the expected one: in left hand trials there was a peak of activity right before the response for the left motor component and a bit later for the left sensory component, while neither component showed significant activation in right hand trials. Those profiles may correspond to the N-10 and N+50 components reported in the literature about motor potentials (Shibasaki and Hallett, 2006).

For comparison, we show activity averages in the same trials for the SCD at C4', the LRP and the motor and sensory linear classifiers. While there are some similarities among them, some differences are evident too. First, the activity picked up by the C4' Laplacian and the LRP filters departed from the profile expected for a purely motor component, since they tended to peak much later, and the period over baseline lasted longer. This probably occurred because the signal they picked up was a mixture of

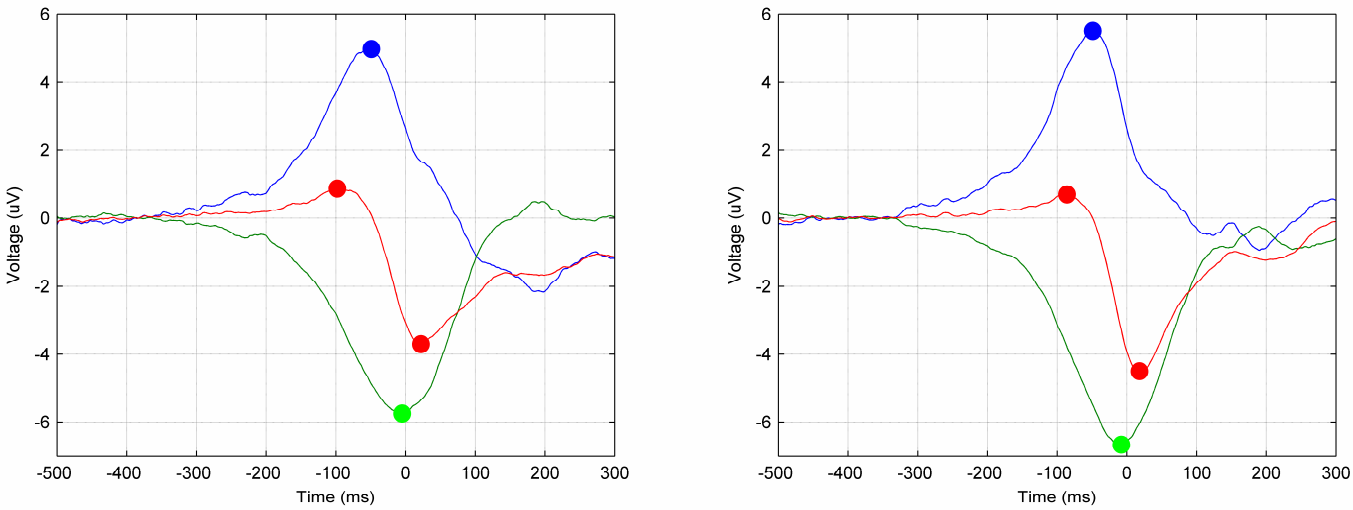

Fig. 5: FCz voltages as reconstructed using only motor components. Left side: left motor components averaged across left hand responses. Right side: right motor components in right hand responses. Blue lines: voltages generated by the motor components. Green lines: voltages generated by the post-movement components. Red line: voltages obtained by addition of motor and post-movement components. The colored circles are located at the peaks of each waveform. The interval between the peak of the motor and post-movement components is much shorter than the distance between the positive and negative peaks of the waveform that results from combining them.

motor and somatosensory potentials. Second, as the filters obtained by multiple regression and logistic regression are tailored to each particular individual, they tend to be less sensitive to artifactual noise than LRP and SCD. The latter filters do not deal with artifactual noise explicitly, while the former are generated with the aim of maximizing discriminability for each subject.

Third, although the sensory linear classifier activity looked reasonable, the activity of the motor classifier also seemed to blend motor and sensory components. This likely happened because the criterion for the generation of the filter was just that it discriminated maximally between left and right hand responses in a certain time window so, given the temporal overlap between motor and somatosensory potentials, the filter could achieve better discriminability by picking up a mixture of

both components. It is hard to overcome this difficulty by merely shifting the time window used for the discrimination. Our method, by contrast, can isolate a component that fits a desired temporal profile while taking into account the variance explained by the rest of the components, which allows to separate the motor contribution from the sensory potentials. Moreover, neither LRP nor classifiers are able, by construction, to dissociate the separate contributions of right and left hand to the EEG potentials.

3.1.2.2 ERP reconstruction and peak latencies

Fig. 5 shows the ERPs for right hand and left hand responses reconstructed using the estimated

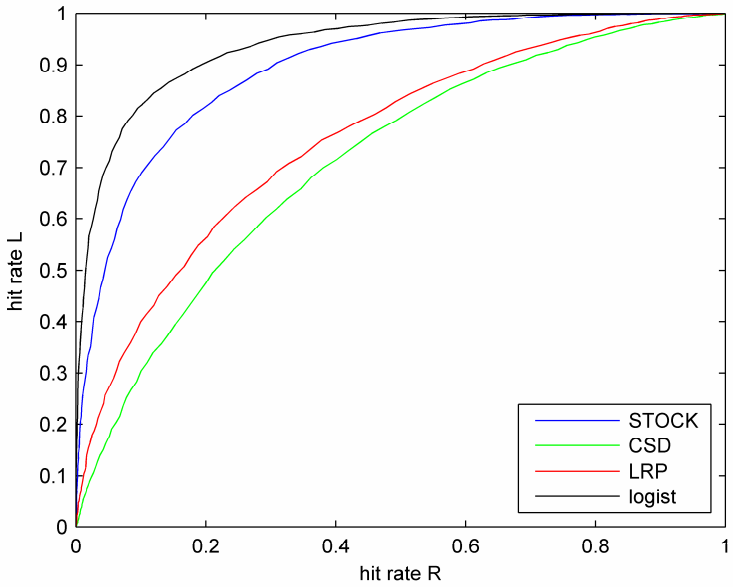

Fig 6. Performance of several spatial filters when discriminating between right and left hand responses, as revealed by the receiver operating characteristic (ROC) curve. The filter for the motor component obtained by logistic regression -as expected by construction- outperforms the combination of the right-hand and left-hand motor filters obtained by the STOCK method, which in turn does better than LRP and than the combination of the C3' and C4' CSD filters. The difference in discrimination performance between STOCK and logistic regression is small, however.

activities of the motor components as obtained by the STOCK method. The interval between the peaks of the isolated pre- and post-movement potentials was 39 ms for left responses and 29 ms for right responses, which match roughly the 35 ms delay we assumed when modeling source activities. Interestingly, this delay is much shorter than the corresponding interval between the maximum and minimum of the waveform obtained combining the two isolated potentials, which was 100 ms for left hand responses, and 102 ms for right responses. This value of around 100 ms, although larger than the 60-70 ms reported in the literature, is congruent with that figure, and validates *ex post* the choice of 35 ms as the interpeak interval in our model, despite its apparent contradiction with ERP data.

3.1.2.3 Left – right discriminability

Left-right discrimination performance for LRP, SCD, linear classifiers and multiple regression is reported in Fig. 6 as the area under the receiver operating characteristic (ROC) curve (Hanley and McNeil, 1982; Parra et al., 2002). Multiple regression outperformed SCD and LRP for all participants, and was surpassed by linear classifiers for all participants. Given that the linear classifiers were generated so that they maximized the discriminability between left and right hand responses, the fact that they topped all the other filters in this comparison came as no surprise. However, there was only a

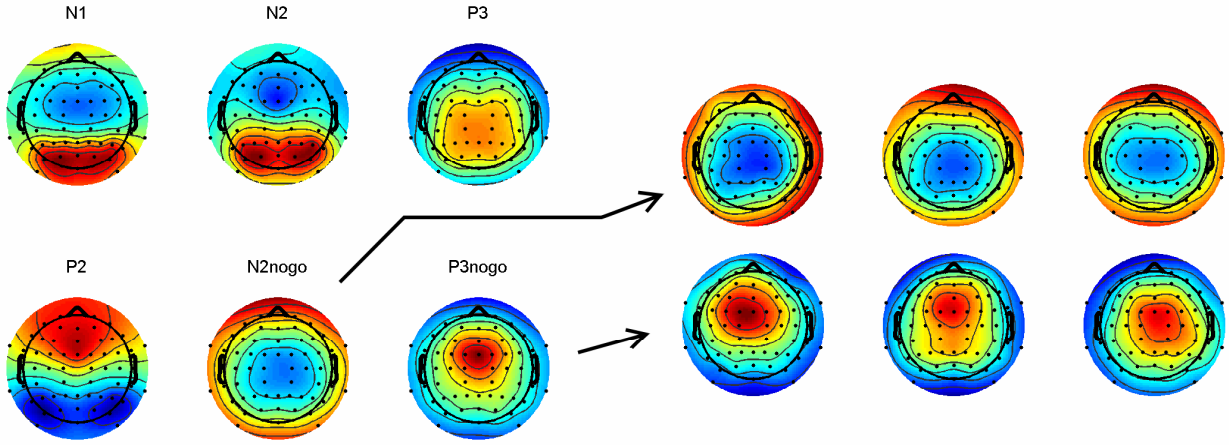

Fig 7. Scalp maps of the perceptual and cognitive components obtained by the STOCK method. Left side: grand average maps. Right side, individual scalp maps for the Nogo-N2 and Nogo-P3 components of three sample subjects. The maps were re-referenced to average offline.

small difference in performance between linear classifiers and the filters for the motor component obtained by STOCK. And, as said before, linear classifiers did not isolate left and right hand motor preparation, and generally failed to separate the motor and somatosensory components. The areas under the average curves were 0.94 for logistic regression, 0.89 for STOCK, 0.76 for LRP and 0.71 for CSD.

3.2 Cognitive N2 and P3 components in dataset 2

3.2.1 Scalp maps

Grand averaged scalp maps for the six modeled components are shown in Fig. 7. For the Nogo-N2 and Nogo-P3, three individual scalp maps from representative subjects are also plotted. The scalp map of the cognitive Nogo-P3 has a fronto-central peak, matching the expected shape for a P3a component (Polich, 2007). In contrast, the scalp map of the generic P3 has a more parietal distribution, more similar to the documented map for the P3b component. The Nogo-N2 has a central peak, slightly posterior to the fronto-central peak that has been reported for the cognitive N2 component. Note, however, that the reported scalp map of the cognitive N2 has been estimated simply by measuring the voltages at a certain time-window, so it may be distorted by the presence of other concurrent event-related potentials. The superposition of the scalp map obtained for the generic N2 and the Nogo-N2 scalp map may well give rise to a scalp voltage distribution similar to that reported for the time-window of the N2 peak. The N1 has a clear occipital map, as expected for a visual perceptual component, while the scalp map of the P2 is more parietal and temporal, with a symmetrical bilateral distribution.

3.2.2 Filters

The no-go ERP waveform obtained after the subtraction of the Nogo-N2 and Nogo-P3 components is shown in Fig. 8 for the FCz channels. As can be seen, removing the Nogo-N2 component makes the modified no-go waveform fit the original go waveform around the N2 peak pretty closely, while maintaining its amplitude and shape for the P3 peak. Conversely, removing the Nogo-P3 component changes the shape of the no-go waveform so that it matches the go waveform around the P3 peak, but keeps it essentially unchanged around the N2 and previous peaks. Accordingly, while the original

amplitude of both N2 and P3 was significantly larger in FCz for correct no-go than for go trials, $t(16) = 3.77$, $p = .002$ in the N2 window (260-320 ms post-stimulus) and $t(16) = 8.44$, $p < .001$, in the P3 window (350-450 ms post-stimulus), subtracting Nogo-N2 and Nogo-P3 from correct no-go trials rendered these differences non-significant in the N2 and the P3 windows, respectively ($t(16) = .34$, $p = .73$; $t(16) = .08$, $p = .93$). However, removing the Nogo-N2 component did not affect the significance of the difference in the P3 window, $t(16) = 8.66$, $p < .001$, nor did removing the Nogo-P3 component reduce the significance of the difference in the N2 window, $t(16) = 4.26$, $p < .001$.

As for the filters obtained using linear discrimination and SOBI, they failed to isolate the cognitive N2 and P3 components selectively. For SOBI, the filter whose output maximally correlated with FCz amplitudes around the N2 peak was also the one that did so around the P3 peak for all participants. The filters obtained by SOBI were thus either sensitive to both the N2 and the P3 simultaneously or to none of them. Subtracting the activity isolated by the N2-P3 SOBI filters reduced drastically the amplitude of both the N2 and P3 deflections in the no-go ERP waveform, to a level similar to that of the go waveform. Although the filters obtained by logistic regression were more selective, subtracting the activity of the putative N2 filter also reduced both N2 and P3 amplitudes, as did removing the activity isolated by the putative P3 filter. Subtracting the activity picked up by the N2 filter from the no-go waveform did cancel the difference between go and correct no-go trials in the N2 time-windows, $t(16) = 1.88$, $p = .08$, but so did removing the putative P3 activity, $t(16) = 1.82$, $p = .09$. Conversely, although

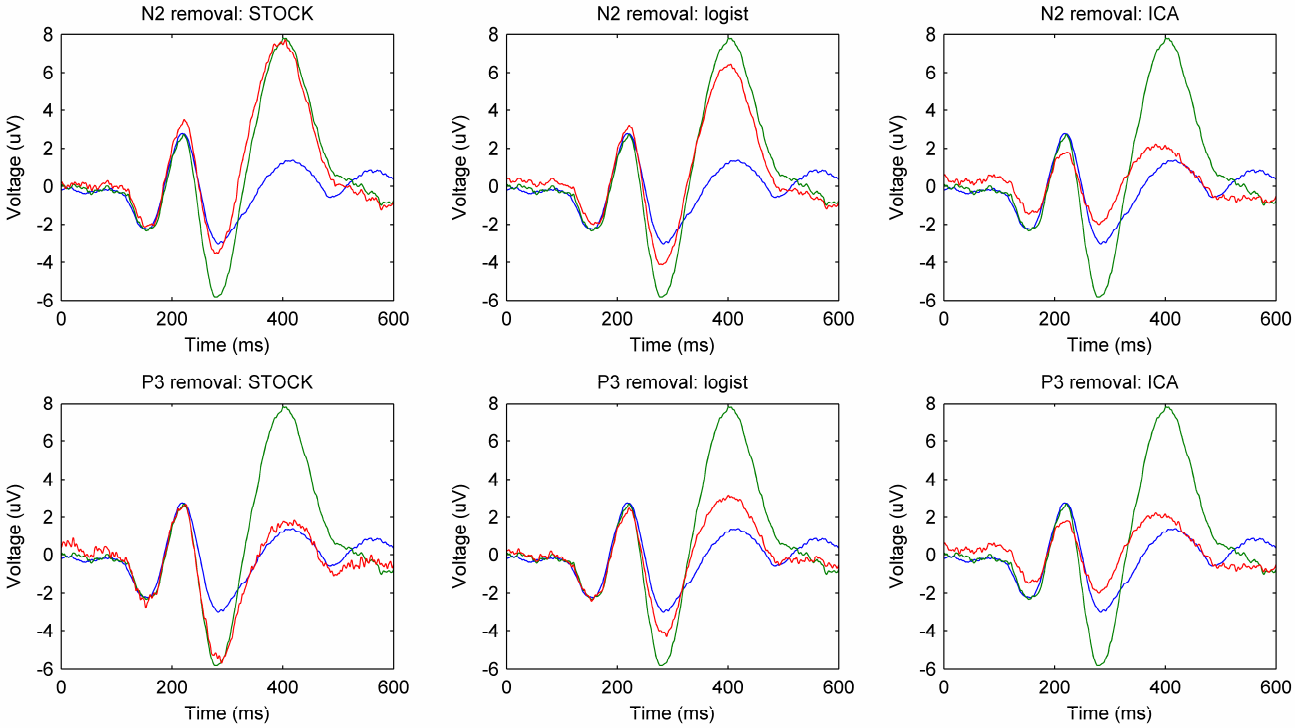

Fig 8. Effect of the subtraction of individual components from the grand average ERP of no-go trials in the FCz channel. The blue and green waveforms are the original ERPs for go and no-go trials respectively, while the red waveform is the grand average of no-go trials after removal of the specified component. Removing the Nogo-N2 component obtained by STOCK rendered the difference between the go and no-go waveforms insignificant in the N2 peak, while leaving unchanged the rest of the waveform. Removing the Nogo-P3 component, conversely, made the no-go waveform match the go trace around the P3 peak, while leaving intact the rest of the ERP. The N2 and P3 filters obtained by logistic regression did not show such selectivity, with the subtraction of any of them affecting both the N2 and the P3 peaks. ICA produced a single filter per participant that was sensitive to N2 and P3 simultaneously.

subtracting the putative N2 activity from correct no-go trials didn't cancel the difference in the P3 window $t(16) = 7.69$, $p < .001$, neither did removing the purported P3 component $t(16) = 3.37$, $p = .004$.

Interestingly, if the activity isolated by the N2 and P3 filters obtained by logistic regression was subtracted not just from no-go trials, but also from go trials, doing so selectively canceled the differences in the corresponding time windows, with $t(16) = .06$, $p = .95$ for the N2 window but $t(16) = 4.40$, $p < .001$ for the P3 window after removing the N2 component, and $t(16) = 8.80$, $p < .001$ for the N2 window but $t(16) = 1.83$, $p = .08$ for the P3 window after removing the P3 component. This suggests that the filters obtained by logistic regression did not actually isolate components that were present in the no-go trials and absent from go trials, as intended. Since these filters just maximize the discriminability between go and correct no-go trials, they capitalize on any difference between them, which may arise from a component that is present in go trials and absent from no-go trials or vice versa. Thus, using linear discrimination there is no way to guarantee that the filters obtained will be selectively sensitive to some component that is present in one particular condition and absent in the other.

Fig. 9 shows the grand average of the estimated activity of the isolated components obtained by

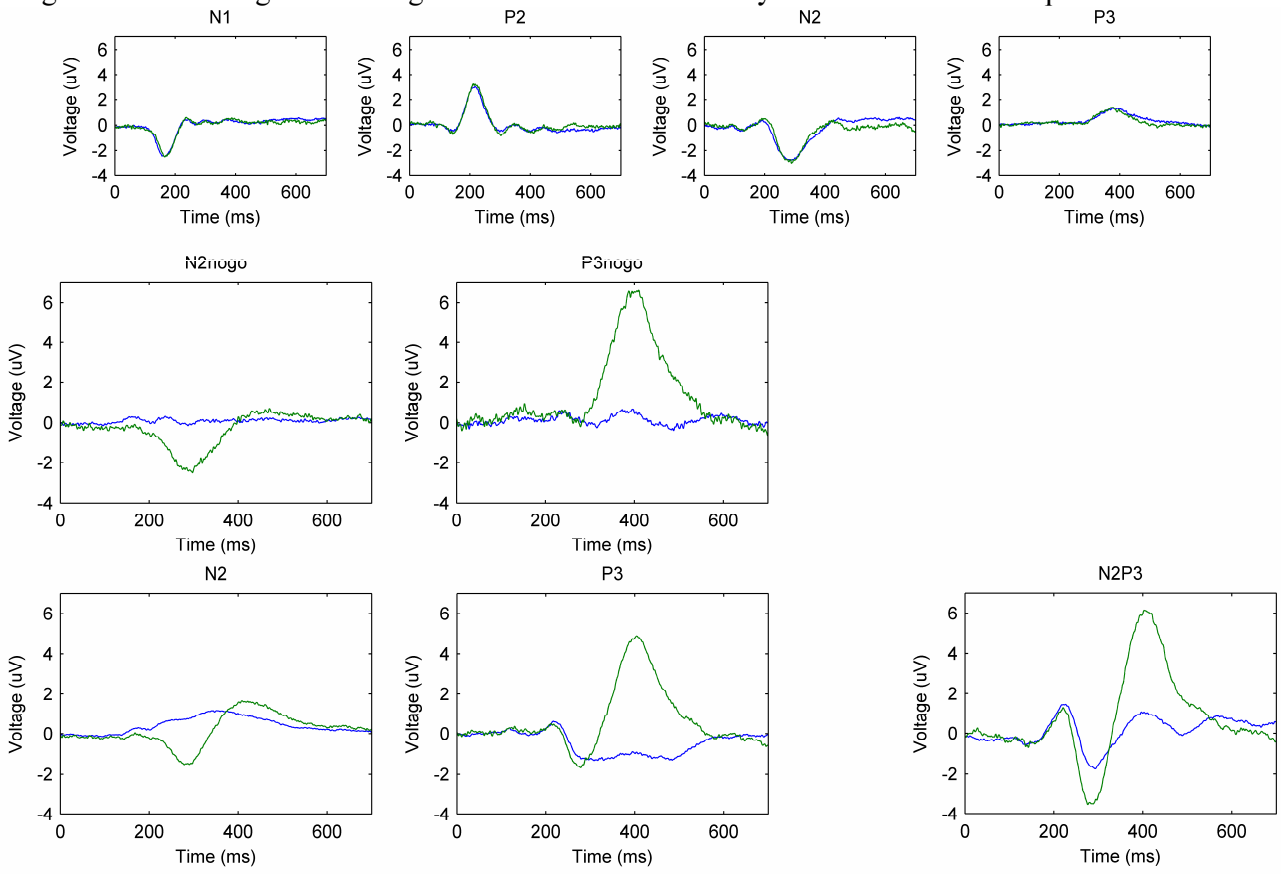

Fig 9. Grand average ERPs of single component activities, projected to FCz. Blue: go trials. Green: no-go trials. Upper row: perceptive and generic components in the STOCK model. Central row: cognitive Nogo-N2 and Nogo-P3 in the STOCK model. Lower row: N2 and P3 filters obtained by logistic regression, and N2P3 filter obtained by ICA (rightmost graph). The early perceptual and generic N2 and P3 components show no difference between go and no-go trials. The cognitive Nogo-N2 and Nogo-P3 components, conversely, display a large activity in no-go trials while keeping almost silent during go trials.

STOCK projected back onto the FCz channel. As expected, the amplitude of the Nogo-N2 and Nogo-P3 components differed significantly between go and correct no-go trials, $t(16) = 5.10$, $p < .001$, $t(16) = 8.61$, $p < .001$, while the rest of the components (the generic N2 and P3 and the N1 and P2 components) showed a similar activity in go and correct no-go trials around their peaks (all p's $> .05$). Note that the curves in the graphs represent an estimation of the electrical activity that could be measured in FCz if only a single component were active each time, which was our original aim.

The combination of the six components obtained by STOCK produces a waveform that matches the original ERP very closely (Fig. 10), meaning that the components included in our model are able to explain almost all event-related variations in amplitude for go/no-go trials.

We were interested in applying the spatial filters obtained from the go/no-go task to the ERPs of the Eriksen task. Fig. 11 shows the grand average of the voltage at FCz for congruent and incongruent trials, after removing the generic N2 and P3 components and after removing the Nogo-N2 and Nogo-P3 components. It is remarkable that, despite the spatial filters for those components having been generated using data from a different task, yet they selectively reduce the amplitude of the expected peaks, i.e., removing the generic N2 component affects mostly the amplitude of the N2 peak, while removing the Nogo-P3 reduces the amplitude of the P3, leaving unchanged those of the other peaks.

Several features of the graphs stand out, though. First, subtracting the Nogo-N2 component leaves the waveforms almost completely intact, while doing so with the generic N2 affects both incongruent and congruent waveforms similarly. This suggests the cognitive N2 component that is present in no-go trials of the go/no-go task might not be the same as the purported cognitive N2 component present in incongruent trials of the Eriksen task. Second, both the generic P3 and the Nogo-P3 components affect the amplitude of the P3 peak, but again, similarly for congruent and incongruent trials. Lastly, the filters are somewhat less selective than they were when applied to the go/no-go trials, with the subtraction of the generic N2 component affecting slightly the amplitudes of the N1 and P3 deflections. This is not surprising since what the proposed method achieves is, in essence, a set of filters that are insensitive to the brain activity and artifactual noise that does not match the temporal profiles of activity proposed in your model. Since these filters were trained with trials of the go/no-go task, they are good at filtering out electrical activity present during that task, but they may fail to do so with other

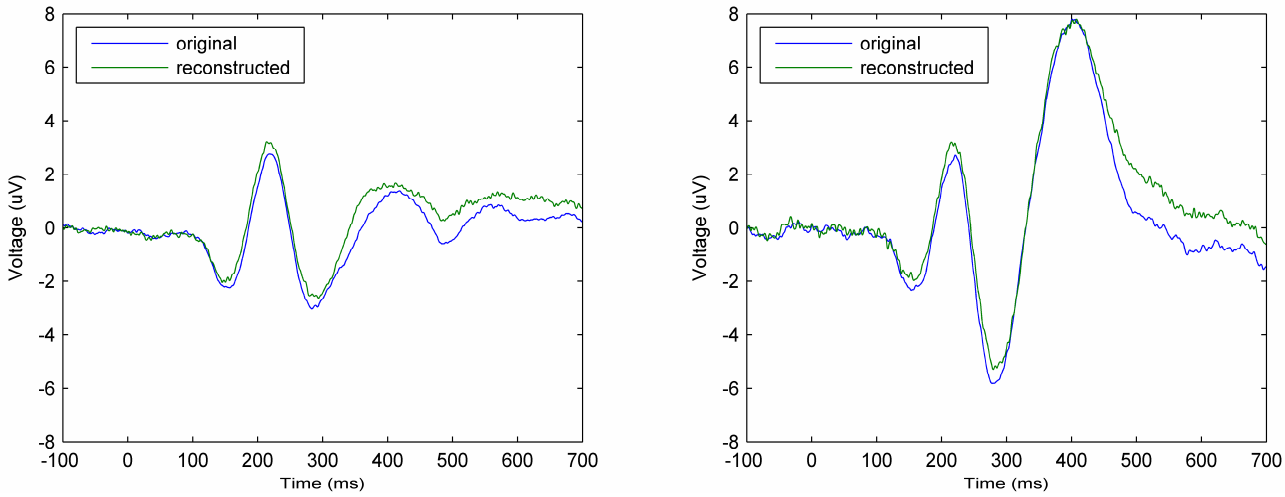

Fig 10. Comparison between original and reconstructed ERPs for go trials (left side) and no-go trials (right side). The reconstructed ERP was obtained by combining the potentials caused by all the six components in the model used for the STOCK method.

brain-related activity that is only present during the flankers task. The fact that, in spite of this, the filters are relatively selective also for the Eriksen task is quite remarkable in our view.

4 Conclusions

Borrowing ideas from the literature on linear classifiers (Parra et al., 2005) and the application of the general linear model to fMRI analysis, we have proposed a method to analyze EEG data by multiple regression. The main contribution of the method is that it makes it possible to tease apart several temporally overlapping ERP components by automatically estimating their relative contribution to the EEG, provided that we have a reasonable hypothesis about their temporal courses. The method is flexible in that allows the researcher to look for components *à la carte*: if there is reason to believe that a component exists having certain temporal profile, the method returns the most likely scalp map for that component and the best possible spatial filter that selectively detects it. Tested on real data, it proved able to isolate components with actual psychological meaning. First, the anatomical source for motor components, as estimated by fitting current dipoles to the scalp maps, was quite close to the brain areas controlling the corresponding psychological processes, while for cognitive components the scalp maps obtained matched closely those reported for the corresponding components by previous conventional ERP studies. Second, the temporal course of the source activities picked up by the filters showed them to be selectively sensitive to the desired components, and the signal-to-noise ratio was better for the filters produced by this method than for other popular spatial filters in use. The method was useful for isolating motor as well as cognitive components. The results seemed to be not very sensitive to the shape and size of the temporal profiles proposed for the psychological processes as long as they were reasonable, which suggests the method may be helpful for isolating components even

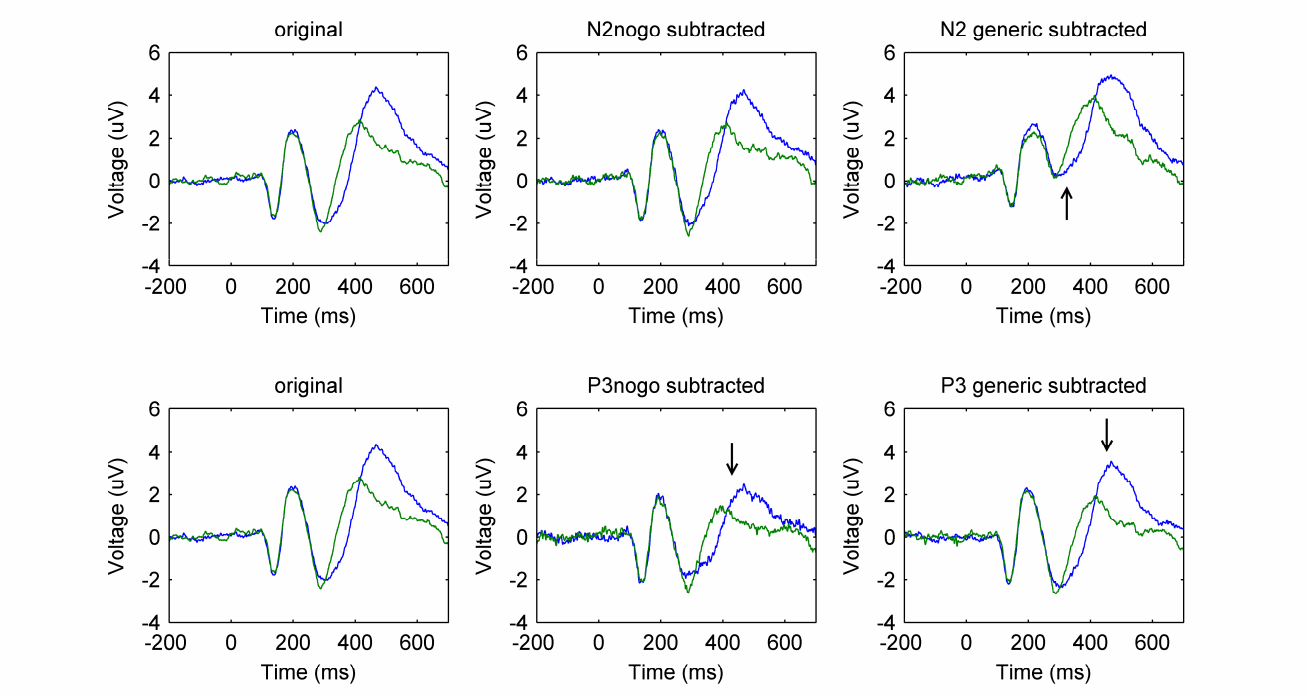

Fig. 11. Comparison between the ERP waveforms in the Eriksen flankers task after subtraction of the generic N2, Nogo-N2, generic P3 and Nogo-P3. Green: congruent trials. Blue: incongruent trials.

when there is some uncertainty about their temporal course.

An arguable limitation of the method proposed here is that it can be very difficult sometimes to generate reasonable hypotheses about the temporal course of the EEG components contributing to the ERPs of a certain task, or even to know how many components are present. There are, however, some cases in which we do have useful information. The literature on local field potentials and unicellular recordings in the cortex of primates contains a wealth of information that can be used to suggest particular shapes for the y_i (see, for instance, the review by Gold and Shadlen (2007) about neural activity related to decision-making). Another rich source of data is the literature about subdural recordings made in humans as a preoperative procedure in epilepsy surgery (Shibasaki and Hallett, 2006). Moreover, the above criticism is equally applicable to logistic regression when used as a method to generate filters selectively sensitive to a particular component: in that case we implicitly assume only that particular component to be more active in one set of trials than in the other. In fact, many, if not most, of the ERP research studies have made similar assumption (Luck, 2005).

The method we propose is an extension to linear classifiers. Classifiers are the optimal choice to build brain-computer interfaces, since the communication there is intended to be digital, not analog. However, the aim of cognitive neuroscientist is to understand brain function, and brain activity evolves gradually. This method enables the researcher to pose gradual hypotheses about the temporal course of psychological processes of interest, something which, in our view, constitutes a refreshing departure from the binary character of most methods of analysis for the EEG. We endorse a comment by Van Berkum, (2008) that "playing 20 Questions with nature (asking simple yes-or-no questions to work something out) is as limited [now] as it was several decades ago, when Alan Newell (1973) made his plea for computational models" (p. 379). For Van Berkum (2008), "The big challenge, then, is to construct precise models that (...) can be mapped onto the neuroimaging data" (p. 379). Though it is true that a binary answer is sometimes enough to reject a theory, we feel that the real advance of neuroscience and cognitive psychology will occur when theories step beyond the realm of qualitative assertions and into the field of quantitative predictions that can be contrasted to measurements. It is our hope that the STOCK method introduced in this paper to isolate gradually changing signals in the EEG helps making easier the mapping from computational models to electrophysiological data.

Acknowledgements

This work has been partly financed by project PSI2012-39292 of the Spanish Ministry of Economy and Competitiveness and a research scholarship granted to D.C. within the III Plan Andaluz de Investigación of the Andalusian Government.

5 References

Albert, J., López-Martín, S., Hinojosa, J.A., Carretié, L., 2013. Spatiotemporal characterization of response inhibition. NeuroImage.

Allison, T., McCARTHY, G., Wood, C.C., Jones, S.J., 1991. Potentials Evoked in Human and Monkey Cerebral Cortex by Stimulation of the Median Nerve a Review of Scalp and Intracranial Recordings. Brain 114, 2465–2503.

Aron, A.R., 2007. The neural basis of inhibition in cognitive control. Neuroscientist 13, 214–228.

Barbati, G., Sigismondi, R., Zappasodi, F., Porcaro, C., Graziadio, S., Valente, G., Balsi, M., Rossini, P.M., Tecchio, F., 2006. Functional source separation from magnetoencephalographic signals. Hum. Brain Mapp. 27, 925–934.

Belouchrani, A., Abed-Meraim, K., Cardoso, J.-F., Moulines, E., 1997. A blind source separation technique using second-order statistics. Signal Process. Ieee Trans. 45, 434–444.

Catena, A., Houghton, G., Valdés, B., Fuentes, L.J., 2009. Unmasking Word Processing with ERPs: Two Novel Linear Techniques for the Estimation of Temporally Overlapped Waveforms. Brain Topogr. 22, 60–71.

Clayson, P.E., Larson, M.J., 2011. Conflict adaptation and sequential trial effects: Support for the conflict monitoring theory. Neuropsychologia 49, 1953–1961.

Coles, M.G., Gratton, G., Donchin, E., 1988. Detecting early communication: Using measures of movement-related potentials to illuminate human information processing. Biol. Psychol. 26, 69–89.

Champod, A.S., Petrides, M., 2010. Dissociation within the frontoparietal network in verbal working memory: A parametric functional magnetic resonance imaging study. J. Neurosci. 30, 3849–3856.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.

Delorme, A., Sejnowski, T., Makeig, S., 2007. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. Neuroimage 34, 1443.

Donders, F.C., 1969. On the speed of mental processes. Acta Psychol. (Amst.) 30, 412.

Durka, P.J., Matysiak, A., Montes, E.M., Sosa, P.V., Blinowska, K.J., 2005. Multichannel matching pursuit and EEG inverse solutions. J. Neurosci. Methods 148, 49–59.

Eimer, M., 1998. The lateralized readiness potential as an on-line measure of central response activation processes. Behav. Res. Methods Instruments Comput. 30, 146–156.

Enriquez-Geppert, S., Konrad, C., Pantev, C., Huster, R.J., 2010. Conflict and inhibition differentially affect the N200/P300 complex in a combined go/nogo and stop-signal task. Neuroimage 51, 877–887.

Falkenstein, M., Hoormann, J., Hohnsbein, J., 1999. ERP components in Go/Nogo tasks and their relation to inhibition. Acta Psychol. (Amst.) 101, 267–291.

Folstein, J.R., Van Petten, C., 2008. Influence of cognitive control and mismatch on the N2 component of the ERP: a review. Psychophysiology 45, 152–170.

Friston, K.J., 1998. Modes or models: a critique on independent component analysis for fMRI. Trends Cogn. Sci. 2, 373–375.

Gajewski, P.D., Falkenstein, M., 2012. Effects of task complexity on ERP components in Go/Nogo tasks. Int. J. Psychophysiol.

Gehring, W.J., Gratton, G., Coles, M.G., Donchin, E., 1992. Probability effects on stimulus evaluation and response processes. J. Exp. Psychol. Hum. Percept. Perform. 18, 198–216.

Gerloff, C., Uenishi, N., Hallett, M., 1998. Cortical Activation During Fast Repetitive Finger Movements in Humans: Dipole Sources of Steady-State Movement-Related Cortical Potentials.

J. Clin. Neurophysiol. 15, 502–513.
Gerloff, C., Uenishi, N., Nagamine, T., Kunieda, T., Hallett, M., Shibasaki, H., 1998. Cortical activation during fast repetitive finger movements in humans: steady-state movement-related magnetic fields and their cortical generators. Electroencephalogr. Clin. Neurophysiol. Mot. Control 109, 444–453.
Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision making. Annu Rev Neurosci 30, 535–574.
Halgren, E., Marinkovic, K., Chauvel, P., 1998. Generators of the late cognitive potentials in auditory and visual oddball tasks. Electroencephalogr. Clin. Neurophysiol. 106, 156–164.
Hanley, J.A., McNeil, B.J., 1982. Maximum attainable discrimination and the utilization of radiologic examinations. J. Chronic Dis. 35, 601–611.
Jung, T.-P., Makeig, S., McKeown, M.J., Bell, A.J., Lee, T.-W., Sejnowski, T.J., 2001. Imaging brain dynamics using independent component analysis. Proc. Ieee 89, 1107–1122.
Karch, S., Feuerecker, R., Leicht, G., Meindl, T., Hantschk, I., Kirsch, V., Ertl, M., Lutz, J., Pogarell, O., Mulert, C., 2010. Separating distinct aspects of the voluntary selection between response alternatives: N2-and P3-related BOLD responses. Neuroimage 51, 356–364.
Knight, R.T., 1996. Contribution of human hippocampal region to novelty detection. Nature 383, 256–259.
Kok, A., Ramautar, J.R., De Ruiter, M.B., Band, G.P., Ridderinkhof, K.R., 2004. ERP components associated with successful and unsuccessful stopping in a stop□signal task. Psychophysiology 41, 9–20.
Kornhuber, Deecke, 1964. Hirnpotentialänderungen beim Menschen vor und nach Willkü rbewegungen, dargestellt durch Magnetbandspeicherung und Rückwärtsanalyse. Pflügers Arch Eur J Physiol 281, 52.
Lee, M.-L.T., Rosner, B.A., 2001. The average area under correlated receiver operating characteristic curves: a nonparametric approach based on generalized two-sample Wilcoxon statistics. J. R. Stat. Soc. Ser. C Appl. Stat. 50, 337–344.
Lee, T.-W., Girolami, M., Sejnowski, T.J., 1999. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. Neural Comput. 11, 417–441.
Løvstad, M., Funderud, I., Lindgren, M., Endestad, T., Due-Tønnessen, P., Meling, T., Voytek, B., Knight, R.T., Solbakk, A.-K., 2012. Contribution of subregions of human frontal cortex to novelty processing. J. Cogn. Neurosci. 24, 378–395.
Luck, S.J., 2005. An introduction to the event-related potential technique. The MIT Press.
Macmillan, N.A., Kaplan, H.L., 1985. Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. Psychol. Bull. 98, 185–199.
Makeig, S., Bell, A.J., Jung, T.-P., Sejnowski, T.J., 1996. Independent component analysis of electroencephalographic data. Adv. Neural Inf. Process. Syst. 145–151.
Nieuwenhuis, S., Aston-Jones, G., Cohen, J.D., 2005. Decision making, the P3, and the locus coeruleus--norepinephrine system. Psychol. Bull. 131, 510.
Onton, J., Delorme, A., Makeig, S., 2005. Frontal midline EEG dynamics during working memory. Neuroimage 27, 341–356.
Onton, J., Westerfield, M., Townsend, J., Makeig, S., 2006. Imaging human EEG dynamics using independent component analysis. Neurosci. Biobehav. Rev. 30, 808–822.
Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 1.
Parra, L., Alvino, C., Tang, A., Pearlmutter, B., Yeung, N., Osman, A., Sajda, P., 2002. Linear spatial integration for single-trial detection in encephalography. NeuroImage 17, 223–230.
Parra, L.C., Spence, C.D., Gerson, A.D., Sajda, P., 2005. Recipes for the linear analysis of EEG.

Neuroimage 28, 326–341.
Pascual-Leone, A., Valls-Solé, J., Wassermann, E.M., Hallett, M., 1994. Responses to rapid-rate transcranial magnetic stimulation of the human motor cortex. Brain 117, 847–858.
Pernier, J., Perrin, F., Bertrand, O., 1988. Scalp current density fields: concept and properties. Electroencephalogr. Clin. Neurophysiol. 69, 385–389.
Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. Electroencephalogr. Clin. Neurophysiol. 72, 184–187.
Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1990. Spherical splines for scalp potential and current density mapping (corrigenda). Electroencephalogr. Clin. Neurophysiol. 76, 565–566.
Pfefferbaum, A., Ford, J.M., Weller, B.J., Kopell, B.S., 1985. ERPs to response production and inhibition. Electroencephalogr. Clin. Neurophysiol. 60, 423–434.
Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. 118, 2128.
Poolman, P., Frank, R.M., Luu, P., Pederson, S.M., Tucker, D.M., 2008. A single-trial analytic framework for EEG analysis and its application to target detection and classification. Neuroimage 42, 787.
Porcaro, C., Coppola, G., Di Lorenzo, G., Zappasodi, F., Siracusano, A., Pierelli, F., Rossini, P.M., Tecchio, F., Seri, S., 2009. Hand somatosensory subcortical and cortical sources assessed by functional source separation: an EEG study. Hum. Brain Mapp. 30, 660–674.
Ramautar, J.R., Kok, A., Ridderinkhof, K.R., 2004. Effects of stop-signal probability in the stop-signal paradigm: the N2/P3 complex further validated. Brain Cogn. 56, 234–252.
Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., Nieuwenhuis, S., 2004. The role of the medial frontal cortex in cognitive control. Sci. Signal. 306, 443.
Rossini, P.M., Martino, G., Narici, L., Pasquarelli, A., Peresson, M., Pizzella, V., Tecchio, F., Torrioli, G., Romani, G.L., 1994. Short-term brain “plasticity”in humans: transient finger representation changes in sensory cortex somatotopy following ischemic anesthesia. Brain Res. 642, 169–177.
Sams, M., Alho, K., Näätänen, R., 1983. Sequential effects on the ERP in discriminating two stimuli. Biol. Psychol. 17, 41–58.
Shibasaki, H., 2012. Cortical activities associated with voluntary movements and involuntary movements. Clin. Neurophysiol. 123, 229–243.
Shibasaki, H., Hallett, M., 2006. What is the Bereitschaftspotential? Clin. Neurophysiol. 117, 2341–2356.
Soltani, M., Knight, R.T., 2000. Neural origins of the P300. Crit. Rev. Neurobiol. 14.
Stone, J.V., 2002. Independent component analysis: an introduction. Trends Cogn. Sci. 6, 59–64.
Tang, A.C., Liu, J.-Y., Sutherland, M.T., 2005. Recovery of correlated neuronal sources from EEG: The good and bad ways of using SOBI. Neuroimage 28, 507–519.
Taniguchi, Y., Burle, B., Vidal, F., Bonnet, M., 2001. Deficit in motor cortical activity for simultaneous bimanual responses. Exp. Brain Res. 137, 259–268.
Van Berkum, J.J., 2008. Understanding Sentences in Context What Brain Waves Can Tell Us. Curr. Dir. Psychol. Sci. 17, 376–380.
Yeung, N., Nieuwenhuis, S., 2009. Dissociating response conflict and error likelihood in anterior cingulate cortex. J. Neurosci. 29, 14506–14510.
Zibulevsky, M., Pearlmutter, B.A., 2001. Blind source separation by sparse decomposition in a signal dictionary. Neural Comput. 13, 863–882.
Zipp, P., 1982. Recommendations for the standardization of lead positions in surface electromyography. Eur. J. Appl. Physiol. 50, 41–54.

Discusión general

A lo largo de las siguientes páginas presentamos primero un resumen general de los resultados obtenidos en la presente tesis doctoral y después los ponemos en el contexto de la investigación actual para completar su significación y alcance teórico, metodológico y práctico.

Emoción e impulsividad

Los resultados del experimento descrito en el capítulo 2 son un indicio claro de que el procesamiento emocional puede provocar un incremento en la urgencia en la toma de decisiones, independiente de los ya conocidos efectos atencionales provocados por la emoción, y más allá de la mera preactivación de determinados programas motores compatibles con el valor afectivo de los estimulos emocionales empleados. Nuestra aportación es importante porque permite diferenciar entre (a) un efecto de interferencia transitorio que comienza de modo inmediato (sólo 200 ms tras la presentación del estímulo) muy probablemente de naturaleza atencional, y más concretamente relacionado con el parpadeo atencional inducido por estímulos emocionales (Dux & Marois, 2009; Smith, Most, Newsome, & Zald, 2006), y (b) un efecto de facilitación que resulta apreciable a 600 ms de la aparición del estímulo, que no puede explicarse en términos atencionales, aunque quizá sí de potenciación perceptiva (Liu, Keil, & Ding, 2012; Schupp, Markus, Weike, & Hamm, 2003), y que es compatible con un incremento en la urgencia en la toma de decisiones. Este segundo efecto facilitador no se había descrito previamente, y es independiente de la respuesta demandada, por lo que es improbable que se deba a que la emoción privilegia en concreto el programa motor utilizado para responder.

Estos resultados, interesantes como análisis de la microestructura de los efectos provocados por el procesamiento emocional, se complementan con los reportados en el capítulo 3, en el que el énfasis se centró más en las consecuencias para la vida real que pueden tener los efectos del procesamiento emocional sobre la conducta de riesgo. Las propias condiciones del experimento, realizado en un escenario simulado de media fidelidad, hacen mucho más difícil el control de las variables extrañas y la interpretación de los resultados obtenidos (Boyle & Lee, 2010; Underwood, Crundall, & Chapman, 2011). Pero, por otra parte, la mayor validez ecológica del experimento hace que éstos tengan mucha más relevancia práctica. En el capítulo 3 se adelantan algunas interpretaciones posibles de los resultados obtenidos como consecuencia de efectos atencionales o motores de la emoción. Además, nuestros datos pueden explicarse también como el resultado de efectos de la emoción sobre la impulsividad. Dado que la variable observada es meramente la mayor propensión del conductor a sufrir un accidente, y la tendencia a mantener la velocidad a pesar de existir una

situación de riesgo en el entorno, bien podría ocurrir que el efecto de la emoción consista en limitar la capacidad del motorista para representar las consecuencias a largo plazo de la conducta arriesgada, que es probablemente el mismo fenómeno que subyace a la incapacidad para soportar la demora de un reforzador. La activación emocional del conductor podría provocar en éste una propensión a valorar más el disfrute inmediato de la conducción a gran velocidad que la posibilidad de sufrir un accidente. Del mismo modo, una interferencia emocional durante la conducción podría dificultar la evaluación del riesgo, entendido como incertidumbre, presente en la situación, del mismo modo que las lesiones ventromediales interfieren con la preferencia de las personas por opciones inciertas (Jones et al., 2012, Capítulo 1).

Medidas fisiológicas del control cognitivo

A pesar del consenso existente sobre la vinculación del componente N2 con el control cognitivo, sigue sin estar clara la función específica que desempeña el proceso que da origen a dicha componente. Las diversas teorías que se han propuesto sobre la función cerebral que refleja –la detección de errores y la evaluación de los propios actos (Alexander & Brown, 2011; Brown & Braver, 2005), la detección de conflictos entre respuestas (Botvinick, Cohen, & Carter, 2004), o la inhibición de respuestas motoras en curso (Jodo & Kayama, 1992; Pfefferbaum, Ford, Weller, & Kopell, 1985)– cuentan todas ellas con evidencia científica que las apoya, y no existe, hasta la fecha, un experimento concluyente que haya permitido descartar ninguna de ellas.

Sin embargo, hay un aspecto del control cognitivo cuya relación con el N2 es susceptible de contrastarse empíricamente de forma relativamente sencilla puesto que tiene un correlato directamente observable: la capacidad de interrumpir acciones incorrectas ya iniciadas (el control de la impulsividad motora, Bokura, Yamaguchi, & Kobayashi, 2001; Donkers & van Boxtel, 2004). El N2 se ha vinculado a esta función, bien como un índice directo de la misma, o indirectamente como un reflejo de la detección de conflictos entre respuestas, que a su vez determina la inhibición de respuestas. Sin embargo, si el N2 está implicado en la cancelación de acciones en curso, su aparición debería preceder a dicha cancelación. Hasta donde sabemos, ningún estudio previo había intentado responder a esta sencilla pregunta, que tiene profundas implicaciones para los diversos modelos teóricos y computacionales que se han propuesto para la función de la corteza prefrontal medial.

En el estudio descrito en el capítulo 4, registramos simultáneamente actividad muscular (EMG) y cerebral (EEG) y empleamos un método automático para determinar, sujeto a sujeto y ensayo a ensayo, la latencia entre la respuesta muscular cancelada y el comienzo del N2, para determinar si el N2 precede a la cancelación de la respuesta o viceversa. Los resultados indican que la cancelación de la respuesta motora empieza bastante antes del comienzo del N2.

Esto lleva a concluir que el N2, así como cualquier otro componente cognitivo posterior, como el P3, no pueden reflejar la actividad del sistema top-down que inhibe las respuestas motoras erróneas ya iniciadas. Su papel debe ceñirse, en todo caso, al de informar o controlar el ajuste de la estrategia de conducta entre ensayos. Lo que nuestros resultados aclaran es que la actividad cerebral reflejada en el N2 es, como mínimo, innecesaria para la inhibición motora, puesto que la inhibición comienza con anterioridad (sin descartar que, de algún modo, esa actividad cerebral pueda contribuir a la misma). Por lo tanto, también previenen contra la interpretación del N2 como el correlato fisiológico de la actividad del mencionado sistema top-down inhibitorio. Nuestros resultados, además, imponen condiciones de contorno para otros modelos de la función del N2, tales como la teoría del bucle de conflicto del cíngulo anterior, que aunque no postulan un papel directo del N2 en la inhibición de respuestas ya iniciadas, sí hacen ciertas predicciones sobre su curso temporal que son difíciles de conciliar con nuestras observaciones.

Las conclusiones del capítulo 4, esto es, la aparente inexistencia de índices en el EEG que reflejen la actividad del sistema inhibitorio top-down (de cuya existencia, por otra parte, se posee evidencia suficiente (Aron & Poldrack, 2006) plantean la pregunta de si es posible desarrollar algún método para aislar en el EEG la actividad originada por un determinado proceso cerebral en el que tengamos interés. El estudio presentado en el capítulo 5 prueba que es posible generar filtros a medida para aislar en el EEG la actividad de procesos cerebrales que se solapan en el tiempo, siempre y cuando se tengan hipótesis correctas sobre el curso temporal de estos procesos. Los filtros obtenidos mediante nuestro método aíslan señales que realmente se comportan como cabe esperar para los procesos cerebrales que se pretende aislar, más allá de la mera reproducción del perfil temporal propuesto para su actividad. Por ejemplo, existen circunstancias en las que podríamos esperar activación y desactivación motora que no se consideraron en las hipótesis sencillas que formulamos sobre la actividad de los procesos motores y somatosensoriales asociados. Para generar los filtros motores supusimos que sólo había actividad motora *antes* de cada respuesta de los participantes. Sin embargo, cuando se pulsa la tecla equivocada en un ensayo de una tarea experimental, existe una tendencia natural a intentar corregir el error pulsando la tecla correcta *después* de haber pulsado la incorrecta. Esto se pudo confirmar en nuestro caso observando el registro electromiográfico de la mano correspondiente a la respuesta correcta, que mostraba actividad después de los errores de pulsación de la tecla equivocada. Pues bien, aunque esta cuestión no se comenta en el capítulo 5, los filtros EEG generados por nuestro procedimiento efectivamente registraban la actividad cerebral correspondiente a los intentos de rectificación, a pesar de que no se modelaron explícitamente en las hipótesis empleadas para generar los filtros.

De manera similar, en nuestro modelo no incluimos explícitamente la inhibición lateral. Durante la selección de respuestas en la tarea de flancos de Eriksen, la competición entre respuestas, y por consiguiente, la inhibición lateral, debe ser mayor en los ensayos incongruentes que en los congruentes. Por tanto, la actividad motora de la mano errónea debería ser inicialmente mayor en los ensayos incongruentes que en los congruentes pero, cuando el conflicto se resuelve, justo antes de responder, dicha actividad debería ser *menor* en los ensayos incongruentes (debido a una inhibición lateral mayor) que en los congruentes (donde la inhibición lateral no es tan intensa). Esto es, efectivamente, lo que se observó al comparar la señal recogida por los filtros de las componentes motoras en los ensayos congruentes e incongruentes, aunque estos resultados se omitieron en la versión final del capítulo.

Por último, nuestro método tampoco produce filtros capaces de generar una señal con cualquier perfil temporal totalmente arbitrario que se proponga. En diversas pruebas realizadas que, por razones de espacio, tampoco se recogen en la versión final del artículo enviado, pudimos comprobar que la generación de filtros sólo funciona cuando existe verdaderamente en el EEG una componente con un perfil de actividad como el hipotetizado. Por otro lado, conocer la forma exacta del perfil temporal tampoco resulta crucial. Ligeras variaciones en la localización temporal y en el perfil de los componentes propuestos tienden a producir filtros y mapas similares, dado que el procedimiento de generación es estadístico, por lo que no requiere una coincidencia exacta entre el verdadero perfil temporal de actividad de las componentes que se desea aislar y el perfil hipotético utilizado para guiar la separación de las mismas.

Conclusión

Nuestro trabajo aporta pruebas de que la emoción puede afectar a la impulsividad, entendida como urgencia en la toma de decisiones (Capítulo 2) o como propensión al riesgo (Capítulo 3). Este efecto de la emoción no es solamente relevante en el plano teórico, sino que puede conllevar consecuencias graves en escenarios reales, como los simulados de forma ecológica mediante el simulador de motocicleta empleado en el Capítulo 3. La evidencia que presentamos, como casi toda la que depende de datos conductuales, está sujeta a interpretación.

La investigación psicológica en general, pero especialmente la relativa a los procesos mentales ha padecido durante décadas el problema de tener que inferir las características de los procesos internos no observables a partir únicamente del output conductual del organismo, un output que es inherentemente ambiguo. Lo ideal para poder caracterizar los procesos cerebrales subyacentes que, por definición, son internos –como lo son todos los relacionados con el control cognitivo– sería medir directamente la actividad del cerebro, y no únicamente variables del comportamiento del individuo, dado que este comportamiento es el resultado de la influencia de múltiples procesos. Una

respuesta puede ser suprimida, por ejemplo, simplemente porque el programa motor se interrumpe, pero también podría ser cancelada por la intervención de un mecanismo de cancelación activo. Este problema histórico ha mejorado de forma considerablemente gracias a la disponibilidad de acceso a instrumentos que permiten medir correlatos directos de la actividad cerebral de forma no invasiva, tanto eléctricos como metabólicos (EEG y fMRI). Sin embargo, estas técnicas, aun reduciendo el problema de la inferencia no han permitido todavía resolverlo plenamente (Niendam et al., 2012). Por ejemplo, dada la limitada resolución temporal de la resonancia magnética funcional, para poder asociar inequívocamente, en un experimento realizado con esta técnica, una cierta área cerebral a una función como la inhibición activa de respuestas motoras, sería necesario aislar el proceso en cuestión en el diseño experimental, de manera que la única diferencia entre dos condiciones experimentales cualesquiera fuera el grado de inhibición, con el resto de los procesos cerebrales idénticos en ambas condiciones (Braver et al., 1997). De no ser así, cualquier diferencia de activación entre ellas podría atribuirse a las diferencias de actividad de esos otros procesos. Sin embargo, como ya se ha comentado, es prácticamente imposible crear dos condiciones en las que el grado de inhibición motora sea diferente, pero el grado de actividad motora sea idéntico. La situación es aparentemente algo mejor en el caso de la electroencefalografía. Sin embargo, aunque el EEG sí dispone, al menos en principio, de la resolución temporal suficiente, no toda la actividad cerebral se refleja en él. El EEG es, además, inherentemente ambiguo, puesto que la solución al problema inverso (localizar las fuentes) para una configuración de voltajes superficiales dada, no es única (Friston et al., 2008; Pascual-Marqui, 1999).

Dentro de la actividad que sí es observable mediante EEG, el componente N2 parece ser un buen candidato a índice de control cognitivo (Folstein & Van Petten, 2008; Johnstone, Barry, & Clarke, 2012; West, Bailey, Tiernan, Boonsuk, & Gilbert, 2012). Nuestro trabajo aporta, no obstante, pruebas convincentes de que el N2 no refleja la actividad de un hipotético sistema responsable de la cancelación de respuestas erróneas ya iniciadas, sino que probablemente refleja alguna forma de feedback de ejecución que el cerebro emplea para ajustar el nivel de control cognitivo para los ensayos sucesivos (Capítulo 4). Si existe o no una forma de medir directamente la inhibición de respuestas motoras mediante EEG es algo que depende de la anatomía humana (podría ser, de hecho, imposible), pero también del desarrollo de técnicas para aislar señales de interés en el EEG. El método que proponemos en el capítulo 5 para generar filtros a medida para procesos específicos a partir de nuestro conocimiento previo sobre su curso temporal esperado, es un paso en esa dirección. Aunque dicho método requiere conocer a priori el perfil temporal de actividad del proceso que deseamos aislar, esa información puede estar disponible como resultado de estudios previos con registros invasivos. El conocimiento acumulado sobre la actividad neuronal mediante registros unicelulares o de potenciales de campo local podrá servir de ese modo para guiar el

análisis del EEG (algo que hasta la fecha rara vez ha ocurrido), mejorando la especificidad de los filtros utilizados para aislar la señal de interés.

Referencias

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature neuroscience*, *14*(10), 1338-1344.

Aron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus. *The Journal of Neuroscience*, *26*(9), 2424-2433.

Bokura, H., Yamaguchi, S., & Kobayashi, S. (2001). Electrophysiological correlates for response inhibition in a Go/NoGo task. *Clinical Neurophysiology*.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in cognitive sciences*, *8*(12), 539-546.

Boyle, L. N., & Lee, J. D. (2010). Using driving simulators to assess driving safety. *Accident Analysis & Prevention*, *42*(3), 785-787.

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*(5712), 1118-1121.

Donkers, F. C., & van Boxtel, G. J. (2004). The N2 in go/no-go tasks reflects conflict monitoring not response inhibition. *Brain and cognition*, *56*(2), 165-176.

Dux, P. E., & Marois, R. (2009). The attentional blink: A review of data and theory. *Attention, Perception, & Psychophysics*, *71*(8), 1683-1700.

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, *45*(1), 152-170.

Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., … Mattout, J. (2008). Multiple sparse priors for the M/EEG inverse problem. *Neuroimage*, *39*(3), 1104-1120.

Jodo, E., & Kayama, Y. (1992). Relation of a negative ERP component to response inhibition in a Go/No-go task. *Electroencephalography and clinical neurophysiology*, *82*(6), 477-482.

Johnstone, S. J., Barry, R. J., & Clarke, A. R. (2012). Ten years on: A follow-up review of ERP research in attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*.

Jones, J. L., Esber, G. R., McDannald, M. A., Gruber, A. J., Hernandez, A., Mirenzi, A., & Schoenbaum, G. (2012). Orbitofrontal Cortex Supports Behavior and Learning Using Inferred But Not Cached Values. *Science*, *338*(6109), 953-956.

Liu, Y., Keil, A., & Ding, M. (2012). Effects of emotional conditioning on early visual processing: Temporal dynamics revealed by ERP single□trial analysis. *Human brain mapping*, *33*(4), 909-919.

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(2), 241-268.

Pascual-Marqui, R. D. (1999). Review of methods for solving the EEG inverse problem. *International journal of bioelectromagnetism*, *1*(1), 75-86.

Pfefferbaum, A., Ford, J. M., Weller, B. J., & Kopell, B. S. (1985). ERPs to response production and inhibition. *Electroencephalography and clinical neurophysiology*, *60*(5), 423-434.

Schupp, H. T., Markus, J., Weike, A. I., & Hamm, A. O. (2003). Emotional facilitation of sensory processing in the visual cortex. *Psychological science*, *14*(1), 7-13.

Smith, S. D., Most, S. B., Newsome, L. A., & Zald, D. H. (2006). An emotion-induced attentional blink elicited by aversively conditioned stimuli. *Emotion*, *6*(3), 523-527.

Underwood, G., Crundall, D., & Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation research part F: traffic psychology and behaviour*, *14*(6), 435-446.

West, R., Bailey, K., Tiernan, B. N., Boonsuk, W., & Gilbert, S. (2012). The temporal dynamics of medial and lateral frontal neural activity related to proactive cognitive control. *Neuropsychologia*.