TESIS DOCTORAL

MECANISMOS DE CONTROL EMOCIONAL:

ÍNDICES COMPORTAMENTALES Y NEURALES

AFFECTIVE CONTROL MECHANISMS:

BEHAVIORAL AND NEURAL INDEXES

DOCTORANDA

SONIA ALGUACIL SÁNCHEZ

DIRECTORES

MARÍA RUZ CÁMARA Y PÍO TUDELA GARMENDIA

DEPARTAMENTO DE PSICOLOGÍA EXPERIMENTAL

UNIVERSIDAD DE GRANADA

FEBRERO 2016

Editor: Universidad de Granada. Tesis Doctorales
Autora: Sonia Alguacil Sánchez
ISBN: 978-84-9125-889-6
URI: http://hdl.handle.net/10481/43720

Universidad de Granada
Departamento de Psicología Experimental
Campus Universitario de Cartuja, s/n
Telf. +34-958243763 - Fax: +34 -958246239
18071 - Granada (España)

MECANISMOS DE CONTROL EMOCIONAL:

ÍNDICES COMPORTAMENTALES Y NEURALES

Tesis Doctoral presentada por Sonia Alguacil Sánchez en el Departamento de Psicología Experimental para aspirar al grado de Doctora en Psicología, en el Programa de Doctorado en Psicología de la Universidad de Granada. En este trabajo se han respetado las pautas que establece la normativa de la Universidad de Granada para la obtención del título de Doctorado Internacional.

La tesis ha sido realizada bajo la dirección de los profesores María Ruz Cámara y Pío Tudela Garmendia, quienes avalan la calidad de la misma, así como la formación de la doctoranda para aspirar al grado de doctor.

La doctoranda Sonia Alguacil Sánchez y los directores María Ruz Cámara y Pío Tudela Garmendia garantizan, al firmar esta tesis doctoral, que el trabajo ha sido realizado por la doctoranda bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Firmado en Granada, a 8 de Febrero de 2016

La doctoranda: Los directores de la tesis:

Fdo.: Sonia Alguacil Sánchez

Fdo.: María Ruz Cámara y Pío Tudela Garmendia

Los estudios presentados en esta tesis han sido financiados a través de una beca del programa nacional de Formación del Profesorado Universitario (FPU).

Contenidos

Índice de tablas y figuras

<u>Capítulo 1: Introducción</u>

*'Como todo el mundo, solo tengo a mi servicio tres medios para evaluar la existencia humana: el estudio de mí mismo, que es el más difícil y peligroso, pero también el más fecundo de los métodos; la observación de los hombres, que logran casi siempre ocultarnos sus secretos o hacernos creer que los tienen; y los libros, con los errores particulares de la perspectiva que nacen entre sus líneas.' (*Marguerite Yourcenar, Memorias de Adriano, Pocket Edhasa, 1986).

1.1. El control cognitivo como elemento necesario en la acción

Nuestro cerebro es un sistema altamente complejo que ha evolucionado en constante interacción con el medio. Uno de los elementos clave de dicha evolución es la flexibilidad con la que podemos procesar la información sensorial y transferirla a una serie de acciones que den respuestas adaptativas ante los cambios que se suceden en el ambiente (Fuster, 2004; Morton, Ezekiel, & Wilk, 2011; Pearson, Heilbronner, Barack, Hayden, & Platt, 2011; Rougier, Noelle, Braver, Cohen, & O'Reilly, 2005). Entender y explicar cómo se organiza nuestro comportamiento de manera flexible para generar estas acciones, en coordinación con el constante cambio en los eventos y demandas del mundo que nos rodea, ha sido y es uno de los objetivos principales perseguidos tanto desde la psicología como desde la neurociencia cognitiva.

El entorno en el que nos movemos está repleto de eventos de los que necesitamos crear una representación mental y un plan de acción eficiente. Sin embargo, dada la complejidad de estos eventos y la sucesión de cambios en las condiciones ambientales, es frecuente encontrar situaciones donde la información que percibimos, y en la que basamos nuestras acciones, genere varias representaciones de un mismo evento de manera simultánea. Dichas representaciones pueden interferir entre sí a lo largo de diferentes niveles de procesamiento de la información, incluyendo su percepción, categorización o selección de respuesta (Botvinick, Cohen, & Carter, 2004). Esto ocurrirá en cualquier situación en la que las distintas acciones posibles sean activadas por representaciones coexistentes en el sistema de procesamiento (Posner & DiGirolamo, 1998).

Uno de los constructos elementales para poder explicar la organización del comportamiento y su flexibilidad es el del control cognitivo. Este se refiere a la habilidad de regular, coordinar y secuenciar el pensamiento y de comportarse de acuerdo a unas reglas, objetivos, o instrucciones específicos representados internamente, aunque esto suponga actuar en contra de otras respuestas de carácter más reflejo o que se encuentren en fuerte competición con la que finalmente es seleccionada (Botvinick et al., 2004; Braver, 2012; Rougier et al., 2005). Por tanto, el control cognitivo incide en la configuración de nuestro sistema mental, particularmente en la ejecución de tareas que suponen un reto y son poco habituales (Botvinick et al., 2004). A este respecto, se han propuesto dos formas de procesamiento de la información. La primera, de carácter más automático, en la que las asociaciones entre el estímulo y la respuesta se desencadenan de manera involuntaria debido al aprendizaje previo, bien asentado, de dicha relación. En la segunda, en cambio, propia de situaciones en las que las que el aprendizaje previo no es de utilidad, se daría un procesamiento más controlado basado en representaciones de objetivos internos con la intención de evitar las asociaciones empleadas de forma rutinaria e imponer aquellas acordes con el objetivo pertinente en el momento (Norman & Shallice, 1986; Schneider & Shiffrin, 1977).

En un esfuerzo por describir las distintas vías de procesamiento en el control de la acción, Norman y Shallice (1986) propusieron un marco teórico, vigente hoy día, cuyo elemento vertebrador es un conjunto de esquemas mentales asociados a secuencias de acción específicas, bien aprendidas, que se activan cuando se dan las condiciones apropiadas, permitiendo así controlar el curso de las acciones. Por lo tanto, desde este modelo se entiende que mediante el aprendizaje y la práctica continua de una serie de acciones, las circunstancias ambientales activarían una serie de esquemas de acción organizados de manera jerárquica para dar lugar a conductas bien establecidas.

Por otro lado, este modelo también considera aquellas situaciones que requieren de la implementación voluntaria, consciente, de algún tipo de control para evitar la selección de esquemas automáticos inadecuados, y la consecuente comisión de errores en el curso de la acción. Esto ocurre cuando las condiciones estimulares activan esquemas que no son más pertinentes en relación a los objetivos actuales del individuo. Entre estas situaciones destacan aquellas en las que se requiere de planificación y desarrollo de estrategias, toma de decisiones, implementación de respuestas novedosas, resolución de un conflicto tanto entre estímulos como entre respuestas, detección de un error, o sencillamente de la superación de un obstáculo en actividades que presentan alguna característica poco habitual (Norman & Shallice, 1986; Posner & Fan, 2004).

Así, el modelo expuesto por Norman y Shallice (1986) contempla fundamentalmente dos mecanismos de control sobre la acción, representados gráficamente en dos dimensiones diferentes (ver Figura 1). El primero, representado en la dimensión horizontal, sería el denominado *Contention Scheduling* (CS), y el segundo, correspondiente a la dimensión vertical, sería el Sistema Atencional Supervisor (SAS; *Supervisory Attentional System*).

El mecanismo operante en la dimensión horizontal, CS, sesgaría la competición entre esquemas mediante la selección del más adecuado para prevenir el uso compartido de los recursos por varios esquemas al mismo tiempo. Esta selección se daría principalmente en base a la activación-inhibición ejercida lateralmente entre los propios esquemas. Una vez que la activación de un esquema de acción supera el umbral necesario, este es seleccionado por encima de los demás. Dicho nivel de activación está determinado por diferentes factores, como por ejemplo la coincidencia entre los eventos desencadenantes de dos esquemas incompatibles. El valor de activación del esquema depende de varios elementos. Uno de los más críticos es el aprendizaje acumulado con el esquema. A medida que una acción se practica de forma rutinaria (se adquiere destreza en su desempeño) se propicia que los esquemas que la controlan se especialicen más en el uso de las estructuras implicadas en su realización. Esta especialización permite reducir la interferencia entre los recursos destinados a la acción, lo que en consecuencia reduciría la necesidad de inhibir otros esquemas. Entre los factores más importantes involucrados en la preponderancia de un esquema, se encuentran el sesgo generado por el mecanismos de CS, la adecuación de los factores disparadores, el peso específico de otros esquemas y, de forma crucial, la influencia ejercida por el SAS (ver Figura 1).

Cuando las respuestas más habituales, seleccionadas mediante el mecanismo horizontal, no son apropiadas se hace necesaria la intervención vertical del SAS. La intervención de este sistema resuelve la competición alterando el peso de los diferentes esquemas, aunque no de manera directa, mediante la suma de activación o inhibición a los mismos. Así, este recurso permite la adaptación del individuo en múltiples situaciones en un ambiente en constante cambio.

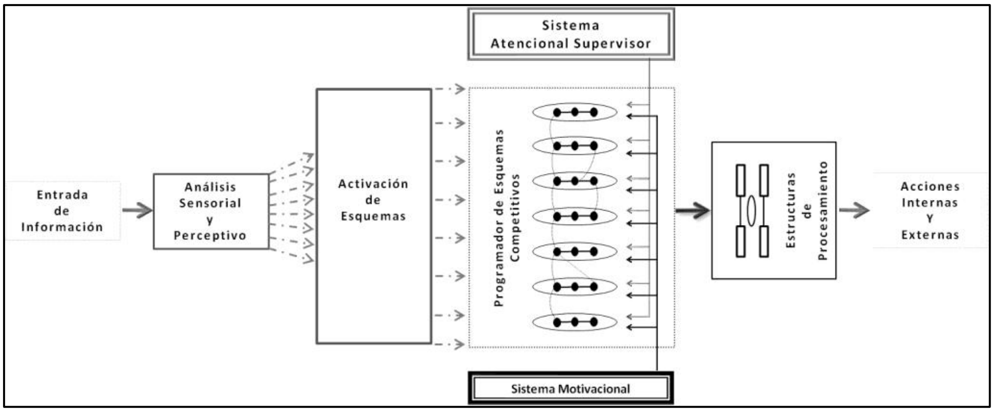

Figura 1: Representación del modelo de Control Ejecutivo de Norman y Shallice.

1.2. La relevancia del conflicto en el control cognitivo

En términos de representaciones mentales y en relación al control cognitivo, el *conflicto* puede ser generado por cualquier composición de estímulos en la que estos o alguna de sus dimensiones generen representaciones asociadas a cursos de procesamiento o acción contrapuestos. Generalmente, uno de ellos es relevante para la resolución de la tarea de forma acorde con los objetivos y otro irrelevante, interfiriendo con el procesamiento del primero y desviando, por tanto, la atención del objetivo (Funes & Lupiáñez, 2003; Kornblum, Hasbroucq, & Osman, 1990).

Bajo este principio se han desarrollado numerosas tareas experimentales, que implican algún tipo de interferencia entre distintos niveles de procesamiento aplicados sobre el estímulo, para el estudio del conflicto y el control cognitivo (Egner, 2008; Funes & Lupiáñez, 2003;

Kornblum et al., 1990). La mayoría de procedimientos empleados caen dentro del marco de los denominados paradigmas de compatibilidad entre estímulo y respuesta, y son aplicables a tareas de tipo espacial, simbólico, sensoriomotor o semántico (Kornblum et al., 1990). Estos paradigmas parten de la idea de que cuando se produce alguna interferencia entre un estímulo y la respuesta es, fundamentalmente, porque en conjunto ambos elementos presentan propiedades comunes. Además, se asume que los estímulos percibidos tienen la capacidad de activar de manera automática respuestas específicas. A este respecto, el modelo integrador realizado por Kornblum y colaboradores (1990) comprende tres posibles combinaciones en función del grado de compatibilidad entre el estímulo y la respuesta: no-compatible, compatible e incompatible (*noncompatible, compatible and incompatible tasks*, denominación extraída originalmente del inglés). El grado de compatibilidad depende de la coincidencia, la similitud y la correspondencia, a todos los niveles, entre el estímulo y la respuesta. Estos criterios son englobados bajo el término de *solapamiento dimensional* (Egner, 2008; Kornblum et al., 1990).

Las tareas dentro de la categoría de no-compatibilidad, son aquellas en las que no hay solapamientos entre las dimensiones y en las que la respuesta activada cumple perfectamente con el objetivo que se persigue. En el caso de las tareas compatibles a nivel de solapamiento dimensional, hay similitud directa entre el estímulo y la respuesta. En consecuencia, como en el caso anterior, la respuesta es fácilmente activada, o inducida, gracias a la relación entre los estímulos. En contraposición a estas dos primeras combinaciones, en las condiciones estimulares para las que hay cierto grado de incompatibilidad, sea cual sea el nivel de procesamiento en el que se produce, la activación de una respuesta se ve obstaculizada.

Las tareas de compatibilidad e incompatibilidad han sido también referidas como situaciones de congruencia e incongruencia, respectivamente. Cuando la situación es de tipo congruente, la resolución del solapamiento ocurre mediante la identificación de la respuesta más conveniente. Sin embargo, cuando hablamos de incongruencia, la resolución de la tarea requiere la creación de una regla, o la búsqueda de una de manera azarosa, para generar la respuesta (Kornblum et al., 1990).

Entre las tareas con cierto grado de solapamiento dimensional más frecuentes en la literatura, se encuentran las de Flancos (Eriksen & Eriksen, 1974), Stroop (Stroop, 1935) y Simon (Simon, 1969), sobre las que se ha efectuado una inmensa variedad de modificaciones (Egner, 2008). Son tareas de elección forzosa en las que la presión temporal es de una importancia crucial para evaluar adecuadamente los mecanismos de control

(Ridderinkhof, 2002). En la Figura 2, quedan recogidas algunas de las versiones más conocidas de las mismas.

Para cada una de las tareas representadas aparecen, de manera esquemática, los estímulos (panel superior), las posibles respuestas (panel intermedio) y las fuentes potencialmente generadoras de conflicto (panel inferior). En cada caso, en la parte inferior se muestra qué tipo de solapamiento dimensional es el responsable de la interferencia (SR, *stimulus relevant dimension*: dimensiones relevantes del estímulo; SI, *stimulus irrelevant dimension*: dimensiones irrelevantes de la tarea; R, *response dimension*: dimensión de respuesta).

Así, por ejemplo, en la tarea Stroop (Stroop, 1935), es necesario indicar el color de la tinta en la que está escrita la palabra (SR) mientras se ignora su significado (SI). Ambas dimensiones coinciden, se solapan, a nivel semántico lo que resulta en dos posibles combinaciones: congruente (color y significado coinciden) e incongruente (cada dimensión señala un color diferente como respuesta posible). En todas ellas, es posible medir, tanto a nivel comportamental como neural, el grado de interferencia o conflicto entre las distintas representaciones. Por ello, estos paradigmas ofrecen herramientas experimentales muy útiles para estudiar tanto la detección como la resolución del conflicto, así como las regiones neurales implicadas en estos procesos, los cuales describiremos en los apartados que siguen.

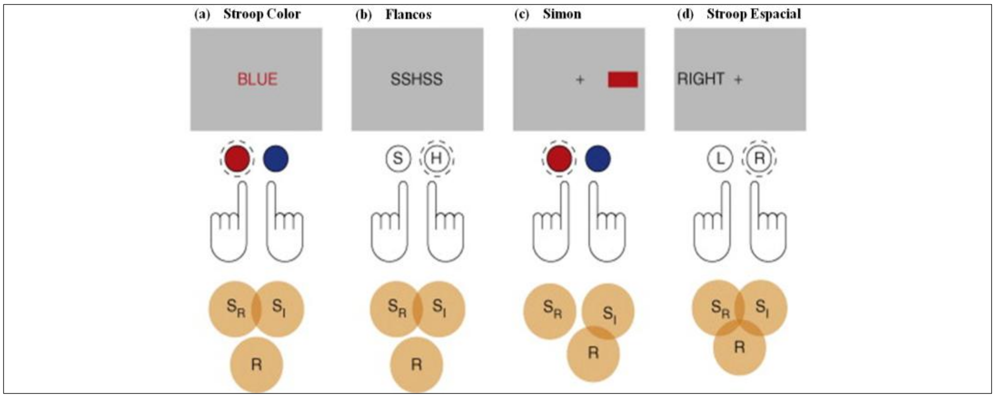

Figura 2: Representación gráfica, tomada de Egner (2008), de las fuentes de conflicto en algunos de los paradigmas de compatibilidad estímulo-respuesta empleados con mayor frecuencia. Stroop (a), Flancos (b), Simon (c) y Stroop espacial (d).

1.3. Efectos comportamentales asociados al conflicto

En todos los casos en los que se produce un solapamiento incongruente, aumentan tanto el tiempo requerido para tomar la decisión como la probabilidad de cometer un error (Egner, 2008; Funes & Lupiáñez, 2003; Kornblum et al., 1990). Esto es lo que se conoce como *efecto de conflicto* (Eriksen & Eriksen, 1974).

La aparición de conflicto, además de entorpecer la acción de forma puntual, genera otro tipo de efectos relacionados. Cuando la selección falla y se comete un error, la respuesta emitida en el ensayo siguiente es más lenta de lo que habría sido si no se hubiese cometido dicho error (i.e. *post-error slowing/post-conflict slowing effect*). Son varias las explicaciones que se han dado a este fenómeno. Por un lado, ha sido atribuido a la puesta en marcha de un modo de control ejercido con un mayor grado de precaución que permitiría reducir la comisión de errores futuros (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Otros autores han propuesto que se debe a la presencia de un procesamiento erróneo de la información que persiste a lo largo de varios ensayos (Gehring, Goss, Coles, Meyer, & Donchin, 1993). Una tercera explicación posible, tal vez la que ha recibido mayor consenso, es que este efecto se debe a una orientación indebida de la atención hacia la aparición de los errores durante la tarea. Dada la infrecuencia de la comisión de errores, estos capturarían la atención generando una orientación anómala cuando aparecen (Notebaert et al., 2009).

Otros factores relacionados con el conflicto son variables moduladoras como el contexto de congruencia o incongruencia previo así como la proporción general de congruencia en la tarea (Botvinick et al., 2001; Logan & Zbrodoff, 1979). Aquí, dada su relevancia para parte de la investigación realizada en la presente tesis doctoral, nos centraremos en los ajustes sobre el efecto de conflicto que tienen lugar como consecuencia de la presencia o ausencia de interferencia en el ensayo antecedente al que se esté resolviendo (Egner, 2008; Gratton, Coles, & Donchin, 1992). La presentación secuencial de dos ensayos con el mismo o distinto tipo de solapamiento dimensional, dependiendo de si la combinación es de tipo congruente o incongruente, puede influir en el grado de conflicto en la segunda presentación. De esta forma, si la sucesión se da entre un ensayo congruente y uno incongruente, se produce una situación de alto conflicto; mientras que si el ensayo incongruente va precedido por otro de la misma naturaleza, se propicia una situación de bajo conflicto. En las situaciones de bajo conflicto, la resolución del solapamiento entre representaciones en el segundo ensayo incongruente es menos costosa en términos atencionales, lo que se refleja en un menor efecto de conflicto. Esta facilitación relativa en la resolución se explica porque los mecanismos de control activados durante el ensayo primero permanecen activos en la situación de conflicto que tiene lugar a continuación y, estando más accesibles, contribuyen a solventar el conflicto de manera más eficiente (Botvinick et al., 2001). En la situación opuesta, de alto conflicto, al permanecer inactivos los mecanismos de control en el ensayo previo, la implementación de estos en el ensayo siguiente resulta más costosa. Esta optimización de las estrategias secuenciales de control ha sido denominada como efecto Gratton (Gratton et al., 1992), o efecto de adaptación al conflicto (Botvinick et al., 2001; Egner, Etkin, Gale, & Hirsch, 2008; Gratton et al., 1992).

1.4. Bases neuronales

1.4.1. El control cognitivo en el contexto de la atención

La habilidad de inhibir respuestas que son inapropiadas, en función del contexto específico en el que se desarrolla nuestra actividad, ha sido clásicamente atribuida al lóbulo frontal (Duncan & Owen, 2000; Koechlin, Ody, & Kouneiher, 2003; Miller & Cohen, 2001; Norman & Shallice, 1986). Esta habilidad de control cognitivo, sin embargo, de forma clásica se enmarca en un conjunto heterogéneo de procesos relacionados con el constructo de la atención. Con posterioridad a la propuesta realizada por Norman y Shallice (1986), se propuso un modelo integrador de la atención (Petersen & Posner, 2012; Posner & Petersen, 1990). Para los autores de este modelo, la atención puede ser entendida como un sistema complejo compuesto por redes especializadas en el que al control sobre los pensamientos y afectos se suman la orientación hacia los eventos percibidos y el mantenimiento del nivel de alerta (Posner & Fan, 2004). Estas funciones corresponderían a la Red de Control Ejecutivo, la de Orientación y la Red de Vigilancia, asociadas cada una a un conjunto de estructuras cerebrales y neuromoduladores químicos bien diferenciados (Posner & Fan, 2004; Rueda, Posner, & Rothbart, 2011) aunque interdependientes en ocasiones (e.g. Callejas, Lupiáñez, & Tudela, 2004; Petersen & Posner, 2012). Las tres redes tendrían un carácter modular, de lo que se desprende que cada una está compuesta por una serie de regiones cerebrales encargadas de realizar los cómputos necesarios para la función específica desempeñada por la red en su conjunto. Más concretamente, la red de Control Ejecutivo, que inicialmente fue simplemente asimilada al SAS de Norman y Shallice (1986; Petersen & Posner, 2012; Posner & Petersen, 1990), actualmente se concibe compuesta por dos redes disociables que operan en distintas escalas temporales. Estas son la Red fronto-parietal y la Red cíngulo-opercular (Dosenbach, Fair, Cohen, Schlaggar, & Petersen, 2008; Petersen & Posner, 2012).

Estas dos redes fueron propuestas por Dosenbach y colaboradores (2008) como dos grandes divisiones integradas en un sistema más complejo. Ese sistema tendría un funcionamiento similar al que se propone desde los modelos de red tipo *small-world*, que permiten explicar las conexiones del sistema nervioso como una red compleja con un alto grado de resiliencia.

Aunque a este tipo de operaciones, *top-down* (arriba-abajo), le han sido atribuidas a un grupo limitado de regiones, como la CPF dorsolateral, la CCA dorsal y la corteza prefrontal superior medial, desde el modelo de las dos redes para el control top-down se resalta la necesidad de incluir un número mayor de regiones, funcionalmente interrelacionadas, que puedan coordinarse para el ajuste de las estrategias de control guiadas por metas

(Dosenbach et al., 2008). De hecho, el modelo separa las tres regiones mencionadas más arriba en dos redes que participarían de manera distinta en la adaptación, ensayo a ensayo, y estabilidad, a lo largo de toda la tarea, de los mecanismos de control.

La Red cíngulo-opercular está conformada por la CPF anterior, la ínsula anterior, la CCA dorsal y el tálamo. Su función principal es la de mantener estable el conjunto de representaciones necesarias durante la ejecución de una tarea. La Red fronto-parietal, por otro lado, está formada por la CPF dorsolateral, el lóbulo parietal inferior, la corteza frontal dorsal, el surco intraparietal, el precúneo y la CCA medial. Su función es la de iniciar y ajustar el control cognitivo, respondiendo de manera diferencial según el *feedback* que recibe. El cerebelo actuaría como una *estación de paso* entre las redes fronto-parietal y cíngulo-opercular, como un mecanismo de detección de errores que contribuiría a la optimización del rendimiento del sistema.

Entre las regiones con un mayor grado de participación en los procesos de control top-down destacan la corteza del cíngulo anterior (CCA) y la corteza prefontal dorsolateral (CPF dorsolateral), coincidiendo, a nivel de citoarquitectura, con las áreas de Brodmann 24/32 y 9/46, respectivamente. Los modelos de control cognitivos centrados en el conflicto proponen funciones distintas para estas dos regiones, las cuales, dada su relevancia, se describen a continuación.

1.4.2. Monitorización y resolución del conflicto

Con el tiempo, diversas aproximaciones han tratado de perfilar la contribución específica de cada una estas estructuras al control cognitivo, ciñéndose principalmente a tres situaciones en las que este es necesario. Estas serían tareas que requieren de la supresión de respuestas dominantes de carácter automático, las que presentan un conjunto de respuestas igualmente válidas y, por tanto, no hay prevalencia de una sobre las demás, requiriendo la selección forzosa de una de ellas y, finalmente, aquellas situaciones en las que hay un alto grado de probabilidades de cometer un error (Botvinick et al., 2004; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999).

Dentro de las teorías sobre la regulación del comportamiento y la cognición, desde la *hipótesis de monitorización del conflicto* se ha planteado la existencia de un sistema integrado por dos componentes: uno encargado de la monitorización constante de la ejecución de la tarea, sensible a la ocurrencia de conflicto, y otro involucrado en la implementación de control y en la selección de una acción adecuada a la situación (Botvinick et al., 2004, 1999; Cohen, Botvinick, & Carter, 2000).

La disociación de las respuestas de la CCA y la CPF dorsolateral entre estos roles se fundamenta en varios estudios, dos de los cuales quedan brevemente recogidos a continuación. En primer lugar, partiendo de la explicación aplicada a la naturaleza del efecto de adaptación al conflicto, Botvinick y colaboradores (2001), propusieron que de estar la CCA más involucrada en la monitorización de la posible aparición de un conflicto que en su resolución, debería activarse de manera más consistente en las situaciones de alto conflicto. Es decir, no sólo ante la ocurrencia de un error, sino también cuando dicha interferencia aparece después de un ensayo de tipo congruente, situación en la que el nivel de conflicto es mayor. En cambio, si la función de la CCA está más relacionada con la estrategias de control, su activación debería ser mayor al darse una situación de bajo conflicto, donde los mecanismos de control están más activos y son más accesibles pero la detección de conflicto es de menor relevancia (Botvinick et al., 2001). Los resultados que obtuvieron mostraron una mayor respuesta de la CCA para las situaciones de alto conflicto, lo que sugiere que esta región es más importante en la detección de interferencias que en la selección y/o implementación de estrategias de control de la acción (Botvinick et al., 2001). La respuesta de la CCA parece, además, ser sensible a la magnitud del conflicto detectado (Etkin, Egner, Peraza, Kandel, & Hirsch, 2006).

De manera complementaria, otro estudio, empleando una versión modificada de la tarea Stroop, combinó dos fases de procesamiento de la información de la tarea con el fin de separar las funciones atribuidas a ambos componentes, la CCA y la CPF dorsolateral. Antes de la presentación del ensayo, a los participantes se les indicaba si debían leer la palabra (lectura, respuesta automática) o nombrar el color (denominación, respuesta controlada) en el que estaba escrita. A continuación, los participantes debían responder ante el estímulo (MacDonald, Cohen, Stenger, & Carter, 2000). Lo que encontraron los autores del estudio es que en la fase de instrucción, donde es necesaria una mayor representación y mantenimiento de las demandas de la tarea, la región más activa era la CPF dorsolateral, cuya activación era mayor con el incremento de la dificultad implicada por las instrucciones. Por otro lado, durante la presentación del estímulo que combinaba color y palabra de tipo congruente o incongruente al que se debía responder, el área con un mayor nivel de respuesta fue la CCA, cuya activación se incrementaba en ensayos incongruentes.

Como consecuencia de los hallazgos de estas y otras investigaciones, la división funcional establecida entre la CCA y la CPF dorsolateral quedó reforzada. Siguiendo dicha división, por tanto, ante la presencia de un posible conflicto durante la ejecución de la tarea, el primer componente del sistema se activaría y se encargaría de reclutar los mecanismos de control necesarios.

Este sistema de monitorización constante realizaría una inspección de los niveles de conflicto presentes en cada situación y posteriormente enviaría una señal de aviso a las estructuras relacionadas con el control (Botvinick et al., 2004, 1999; Cohen et al., 2000). El segundo componente, encargado de la implementación de los ajustes estratégicos sobre los mecanismos de control, ha sido vinculado con la respuesta de la CPF dorsolateral (Botvinick et al., 2001, 1999; MacDonald et al., 2000).

El sesgo ejercido sobre la competición depende en parte de la cooperación entre estructuras, de forma acorde con el modelo de la *competición sesgada* (*biased competition*, término original en inglés; ver Beck & Kastner, 2009; Desimone & Duncan, 1995; Duncan, 1996). Dos principios de esta teoría son de especial relevancia. El primero señala que el control sobre la competición se puede ejercer en cada sistema donde se encuentren las propias representaciones o bien a través de otros sistemas en base a la representación de la meta que se persigue. Así, el sesgo puede, por ejemplo, ser generado por una señal a través de mecanismos de control tipo top-down, en base al conocimiento previo o guiado por metas internamente representadas, o bien por medio de mecanismos de tipo *bottom-up* (abajo-arriba), estando el sesgo en este último caso originado en las primeras fases de percepción del estímulo. El segundo principio hace referencia a la integración de la competición entre varios sistemas. Por ejemplo, la representación del estímulo en uno de los sistemas, generaría que dicho evento tuviese una representación mayor también en el resto de sistemas (Beck & Kastner, 2009).

En base a ambos principios, las estrategias de control permiten el establecimiento de una comunicación funcional entre áreas de control y otras más relacionadas con el procesamiento de la información presentada en la tarea. Esta comunicación debe servir para que las áreas de control coordinen la actividad del resto de estructuras involucradas, incrementando o disminuyendo su respuesta neuronal para así eliminar la confusión entre los distintos sistemas cerebrales (Bush, Luu, & Posner, 2000). La competición entre las representaciones distribuidas en diferentes sistemas depende no sólo del valor mismo del esquema, sino también del valor que tienen el resto de esquemas con los que entran en competición (Norman & Shallice, 1986). Esta interrelación entre estructuras de control y regiones de representación de información específica resalta en el campo de la investigación que estudia cómo el contexto estimular y/o de tarea afecta a la especificidad de los mecanismos de control implicados en la resolución de conflicto, un ámbito esencial para la presente tesis doctoral. A continuación finalizamos este apartado mencionando brevemente algunos estudios relacionados con la comparación de control aplicado sobre estímulos con contenido neutro o emocional.

1.5. Estudios clásicos sobre el control emocional

La CCA no solo está compuesta, a nivel citoarquitectónico, por distintas subdivisiones (por ejemplo, porciones dorsal, rostral y ventral) sino que presenta, además, segregación funcional (Bush et al., 2000; Margulies et al., 2007). La atribución de determinados roles a las distintas porciones de la CCA comenzó en parte con los resultados obtenidos mediante una revisión sistemática, de la literatura sobre conflicto cognitivo y emocional realizada por Bush y colaboradores (2000). Comparando los datos de neuroimagen obtenidos durante la realización de diversas tareas de conflicto, observaron que mientras que la porción dorsal de la CCA responde preferiblemente ante situaciones de interferencia entre estímulos de carácter neutro (conflicto cognitivo), la porción ventral se activa en mayor grado ante situaciones de incongruencia cuando existe información afectiva (conflicto emocional). Ambas subdivisiones parecen ejercer cierto grado de inhibición mutua cuando se producen uno u otro tipo de conflicto. Es decir, cuando se produce una interferencia de tipo cognitivo, mientras que la porción más dorsal responde de manera significativa, la activación en la porción ventral disminuye y viceversa.

Cada una de estas subdivisiones funcionales posee comunicaciones propias con otras regiones para la implementación de los mecanismos de control. La CCA ventral está interconectada con estructuras como la amígdala, el núcleo accumbens, el hipotálamo, la ínsula anterior, el hipocampo y la corteza orbito frontal, además de interactuar con los sistemas endocrino y viscero-motor (Bush et al., 2000; Egner et al., 2008; Ochsner, Hughes, Robertson, Cooper, & Gabrieli, 2009). Ante situaciones de conflicto de tipo cognitivo, la CCA dorsal establece comunicación, entre otras regiones, con la CPF dorsolateral, áreas de la corteza parietal y zonas corticales premotoras (e.g. Bush et al., 2000; Egner & Hirsch, 2005). La comunicación funcional que se establece entre las distintas regiones parece depender de la información relevante que actúa como *target* y del tipo de conflicto que se ha de resolver, cognitivo o emocional. Algunos estudios sugieren que mientras que el conflicto cognitivo se resuelve mediante una facilitación del procesamiento de la información relativa al target, en el caso del conflicto de tipo emocional se produce una inhibición de la información irrelevante para la tarea, por medio del decremento de la activación de regiones que participan directamente en su procesamiento (e.g. Egner & Hirsch, 2005; Etkin et al., 2006).

Tener en cuenta la segregación funcional observada en la CCA es relevante por dos motivos distintos. Primero, por las implicaciones que conlleva a la hora de entender el funcionamiento del mecanismo encargado de la detección del conflicto. Es importante, además, porque supuso un avance de gran utilidad en el estudio comparativo de los mecanismos de control implementados sobre materiales con y sin carga afectiva (afectivos y cognitivos, respectivamente).

Sin embargo, en la mayoría de los estudios acerca del conflicto emocional, se ha omitido la importancia de los factores sociales para los fenómenos afectivos (e.g. Parkinson, 1996). De los pocos intentos realizados hasta la fecha, tan solo algunos estudios han considerado el conflicto afectivo emergente entre distintas claves sociales y el rol de los mecanismos de control para su resolución (e.g. Zaki, Hennigan, Weber, & Ochsner, 2010).

En el capítulo siguiente quedan detallados los principales objetivos que han motivado los estudios realizados en el presente trabajo así como la aproximación experimental empleada.

1.6. Referencias

Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision research*, *49*(10), 1154-1165. http://doi.org/10.1016/j.visres.2008.07.012

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624-652.

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. http://doi.org/10.1016/j.tics.2004.10.003

Botvinick, M. M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*(6758), 179-181. http://doi.org/10.1038/46035

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106-113. http://doi.org/10.1016/j.tics.2011.12.010

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4*(6), 215-222.

Callejas, A., Lupiáñez, J., & Tudela, P. (2004). The three attentional networks: on their independence and interactions. *Brain and Cognition*, *54*(3), 225-227. http://doi.org/10.1016/j.bandc.2004.02.012

Cohen, J. D., Botvinick, M. M., & Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: who's in control? *Nature Neuroscience*, *3*(5), 421-423. http://doi.org/10.1038/74783

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*(1), 193-222. http://doi.org/10.1146/annurev.ne.18.030195.001205

Dosenbach, N. U. F., Fair, D. A., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in cognitive sciences*, *12*(3), 99-105. http://doi.org/10.1016/j.tics.2008.01.001

Duncan, J. (1996). Cooperating brain systems in selective perception and action. *Attention and Performance XVI*, *16*.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*(10), 475-483.

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences*, *12*(10), 374-380. http://doi.org/10.1016/j.tics.2008.07.001

Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2008). Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(6), 1475-1484. http://doi.org/10.1093/cercor/bhm179

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, *8*(12), 1784-1790. http://doi.org/10.1038/nn1594

Eriksen, B. A., & Eriksen, C. W. (1974). Eriksen BA, Eriksen CW. Effects of noise letters upon identification of a target letter in a non-search task. Percept Psychophys 16: 143-149. *Perception & Psychophysics*, *16*(1), 143-149. http://doi.org/10.3758/BF03203267

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron*, *51*(6), 871-882. http://doi.org/10.1016/j.neuron.2006.07.029

Funes, M. J., & Lupiáñez, J. (2003). La teoría atencional de Posner: una tarea para medir las funciones atencionales de Orientación, Alerta y Control Cognitivo y la interacción entre ellas.

Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in Cognitive Sciences*, *8*(4), 143-145. http://doi.org/10.1016/j.tics.2004.02.004

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A Neural System for Error Detection and Compensation. *Psychological Science*, *4*(6), 385-390. http://doi.org/10.1111/j.1467-9280.1993.tb00586.x

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *Journal of Experimental Psychology. General*, *121*(4), 480-506.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The Architecture of Cognitive Control in the Human Prefrontal Cortex. *Science*, *302*(5648), 1181-1185. http://doi.org/10.1126/science.1088545

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility--a model and taxonomy. *Psychological Review*, *97*(2), 253-270.

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, *7*(3), 166-174. http://doi.org/10.3758/BF03197535

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science (New York, N.Y.)*, *288*(5472), 1835-1838.

Margulies, D. S., Kelly, A. M. C., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2007). Mapping the functional connectivity of anterior cingulate cortex. *NeuroImage*, *37*(2), 579-588. http://doi.org/10.1016/j.neuroimage.2007.05.019

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167-202. http://doi.org/10.1146/annurev.neuro.24.1.167

Morton, J. B., Ezekiel, F., & Wilk, H. A. (2011). Cognitive Control: Easy to Identify But Hard to Define. *Topics in Cognitive Science*, *3*(2), 212-216. http://doi.org/10.1111/j.1756-8765.2011.01139.x

Norman, D., & Shallice, T. (1986). Attention to Action: Willed and Automatic Control of Behavior. En R. Davidson, R. Schwartz, & D. Shapiro (Eds.), *Consciousness and Self-Regulation: Advances in Research and Theory IV*. Plenum Press.

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition*, *111*(2), 275-279. http://doi.org/10.1016/j.cognition.2009.02.002

Ochsner, K. N., Hughes, B., Robertson, E. R., Cooper, J. C., & Gabrieli, J. D. E. (2009). Neural systems supporting the control of affective and cognitive conflicts. *Journal of Cognitive Neuroscience*, *21*(9), 1842-1855. http://doi.org/10.1162/jocn.2009.21129

Parkinson, B. (1996). Emotions are social. *British Journal of Psychology (London, England: 1953)*, *87 (Pt 4)*, 663-683.

Pearson, J. M., Heilbronner, S. R., Barack, D. L., Hayden, B. Y., & Platt, M. L. (2011). Posterior cingulate cortex: adapting behavior to a changing world. *Trends in Cognitive Sciences*, *15*(4), 143-151. http://doi.org/10.1016/j.tics.2011.02.002

Petersen, S. E., & Posner, M. I. (2012). The Attention System of the Human Brain: 20 Years After. *Annual review of neuroscience*, *35*, 73-89. http://doi.org/10.1146/annurev-neuro-062111-150525

Posner, M. I., & DiGirolamo, G. J. (1998). Executive Attention: Conflict, Target Detection, and Cognitive Control. *The Attentive Brain*.

Posner, M. I., & Fan, J. (2004). Attention as an Organ System. En J. R. Pomerantz (Ed.), *Topics in Integrative Neuroscience: From Cells to Cognition* (Vol. 9780521869133, pp. 31-61). Carnbridge University Press.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25-42. http://doi.org/10.1146/annurev.ne.13.030190.000325

Ridderinkhof, K. R. (2002). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. *Biological Psychology - BIOL PSYCHOL*, 494-519.

Rougier, N. P., Noelle, D. C., Braver, T. S., Cohen, J. D., & O'Reilly, R. C. (2005). Prefrontal cortex and flexible cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(20), 7338-7343. http://doi.org/10.1073/pnas.0502455102

Rueda, M. R., Posner, M. I., & Rothbart, M. (2011). Attentional control and self-regulation. *Handbook of Self-Regulation: Research, Theory and Applications*, 283-300.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1-66. http://doi.org/10.1037/0033-295X.84.1.1

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, *81*(1), 174-176. http://doi.org/10.1037/h0027448

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643-662. http://doi.org/10.1037/h0054651

Zaki, J., Hennigan, K., Weber, J., & Ochsner, K. N. (2010). Social Cognitive Conflict Resolution: Contributions of Domain-General and Domain-

Specific Neural Systems. *The Journal of Neuroscience*, *30*(25), 8481-8488. http://doi.org/10.1523/JNEUROSCI.0382-10.2010

Capítulo 2: Objetivos

Como avanzamos en el capítulo introductorio, el estudio del control cognitivo ha proporcionado una serie de índices comportamentales y neurales asociados a los mecanismos involucrados en la detección y resolución del conflicto. Estos índices han sido obtenidos principalmente en el marco de los paradigmas de conflicto estímulo-respuesta, y en gran parte de ellos los estímulos son de naturaleza cognitiva, carente de carga afectiva. Sin embargo, esto no atiende a la complejidad de las situaciones cotidianas.

Si bien es cierto que diariamente lidiamos con elementos de nuestro entorno que no presentan contenido emocional alguno, muy a menudo nuestras actividades se desarrollan en contextos que en sí mismos presentan o elicitan eventos afectivos. En consecuencia, el esfuerzo por incluir información afectiva en el estudio de los mecanismos de control no es una cuestión baladí. Dos fundamentos añadidos apremian la inclusión de material afectivo en los paradigmas clásicos de compatibilidad estímulo-respuesta. Por un lado, investigaciones previas atestiguan la influencia que la presentación simultánea de materiales afectivos ejerce durante la resolución del conflicto, dado que pueden alterar la rapidez y la eficacia con la que esta se produce (e.g. Garcia-Garcia, Domínguez-Borràs, SanMiguel, & Escera, 2008; Kanske & Kotz, 2010, 2011; Whalen, Bush, Shin, & Rauch, 2006). De otra parte, numerosas propuestas teóricas abogan por la existencia de una estrecha relación entre las emociones y el desempeño normal de nuestras actividades cotidianas. Muchas de las tareas que realizamos suelen estar contextualizadas por nuestra naturaleza social y la estrecha relación que existe entre esta y las emociones (Barrett, 2012; Barrett & Bliss-Moreau, 2009; Ibanez, Kotz, Barrett, Moll, & Ruz, 2014; Parkinson, 1996; Maria Ruz, Ibanez, Kotz, Barrett, & Moll, 2014). Por consiguiente, una aproximación al estudio de los mecanismos de control en entornos que combinen las emociones y el contexto social es fundamental para comprender qué factores, y cómo, determinan nuestro comportamiento habitual en situaciones de conflicto reales.

Los intentos en pos de equiparar el estudio del conflicto con materiales cognitivos y afectivos han sido más frecuentes mediante el empleo de técnicas de neuroimagen, como la resonancia magnética funcional. No obstante, hay cierto vacío de conocimiento sobre el comportamiento de los índices electrofisiológicos durante la resolución del conflicto de tipo afectivo. La flexibilidad con la que resolvemos las interferencias en el curso del procesamiento de la información es una habilidad considerada fundamental por la rapidez con la que hace posible la acomodación y el reajuste de nuestro sistema cognitivo para dar respuesta a las demandas externas. Por tanto, la cronometría de los mecanismos desplegados en situaciones que requieren de estrategias atencionales de reajuste y control es crítica para su correcto desempeño, relevancia que se extiende a su estudio.

Teniendo en cuenta las consideraciones previas el presente trabajo, a lo largo de cuatro series experimentales, se ha centrado en dar respuesta a dos cuestiones fundamentales. La primera se focaliza en la comparación de los procesos involucrados en la resolución del conflicto cognitivo y emocional. Avanzando sobre estudios anteriores (e.g. Kanske & Kotz, 2010a, 2010b, 2011; Whalen et al., 2006), en nuestro caso el contenido emocional es parte integrante de la tarea que realiza la persona. La segunda atañe al estudio del conflicto, a nivel comportamental y neural, en situaciones en las que las emociones son contextualizadas con un referente social frecuente en nuestras vidas, las relaciones interpersonales.

Las series experimentales que han sido realizadas para profundizar en el estudio de ambas cuestiones para darles respuesta se resumen a continuación.

2.1. Conflicto cognitivo y emocional. Serie Experimental I

Con esta primera serie experimental pretendíamos comparar los mecanismos cerebrales involucrados en la resolución del conflicto cuando tiene lugar entre elementos cognitivos o afectivos. Tras la revisión de los estudios relacionados con nuestro objetivo, decidimos emplear una versión modificada del paradigma empleado por Ochsner y colaboradores (2009). Se trata de una tarea tipo flancos (Eriksen & Eriksen, 1974) que utiliza palabras como estímulos. Dos de estas palabras, idénticas entre sí, denominadas flancos enmarcarían a una tercera, llamada target, que quedaría localizada en el centro de la pantalla. La escasa separación entre estas palabras hace que sean percibidas en bloque. El solapamiento dimensional entre las palabras podía deberse, en la versión cognitiva, al grado de compatibilidad entre las categorías semánticas (animal u objeto doméstico) de pertenencia de flancos y target. En el caso de la versión emocional, el solapamiento estaba determinado por la coincidencia entre la valencia afectiva de las 3 palabras (positiva o negativa). A los participantes se les pedía que indicasen, ensayo a ensayo, la categoría semántica (versión cognitiva) o la valencia afectiva (tarea emocional) del estímulo objetivo.

Para poder evaluar los mecanismos cerebrales subyacentes a la resolución de ambos tipos de conflicto, durante la realización de la tarea se registró la actividad de los participantes mediante electroencefalografía de alta densidad (HDEEG; *high density electroencephalography*). Esta técnica permite el registro continuo, con gran precisión temporal, de las fluctuaciones de voltaje en amplios grupos de neuronas.

Los análisis realizados sobre el registro de la actividad cerebral, se centraron en comparar ambas tareas en las etapas de procesamiento susceptibles de mostrar cambios asociados a la implementación de mecanismos control. Estos incluyeron los potenciales P1, N170, N2 y P3, asociados en la literatura a procesamiento perceptivo, categorial, procesos de control y de decisión, respectivamente.

Una vez extraídos los índices propios de cada tipo de conflicto, así como los comunes entre ambos, en las series experimentales posteriores nos centramos en el estudio de la segunda cuestión fundamental para el presente trabajo.

2.2. Series experimentales centradas en el estudio del conflicto emocional en un contexto social: Series II, III y IV

Una comprensión completa de las estrategias de control durante la resolución del conflicto emocional pasa, necesariamente, por la integración de este en un contexto de tipo social. Esto es así debido a la estrecha relación que existe entre los fenómenos afectivos y la naturaleza social propia del ser humano (Barrett, 2012; Ibanez et al., 2014; Keltner & Gross, 1999; Parkinson, 1996; Ruz et al., 2014). Desde un punto de vista ecológico, las emociones contribuyen a una mejor adaptación al medio dado que permiten un análisis rápido, automático, de las condiciones del contexto que ayudan a seleccionar la estrategia de acción más apropiada (Frijda, 1988; Frijda & Mesquita, 1994). Constituyen, junto con la información personal, una de las claves más informativas durante las interacciones personales. Gracias a ellas nos creamos expectativas acerca del comportamiento de los otros. Estas expectativas sientan la base de cualquier predicción que quisiéramos hacer en relación a las consecuencias de nuestras interacciones sociales, y más concretamente al comportamiento de los otros (Moser, Gaertig, & Ruz, 2014; Oosterhof & Todorov, 2008; María Ruz, Madrid, & Tudela, 2013; Todorov, Mandisodza, Goren, & Hall, 2005; Tortosa, Lupiáñez, & Ruz, 2013; Tortosa, Strizhko, Capizzi, & Ruz, 2013).

Sin embargo, las expectativas generadas a partir de la información personal y las emociones percibidas pueden dar lugar a predicciones contradictorias. De manera similar a lo que ocurre en los paradigmas clásicos de compatibilidad estímulo-respuesta, el grado de congruencia de la información emocional con las expectativas generadas por la identidad de una persona puede provocas una disrupción durante la selección de esquemas de acción. Algunos estudios previos han mostrado que este tipo de conflicto puede observarse en adaptaciones de juegos de tipo económico (María Ruz et al., 2013; María Ruz & Tudela, 2011; Tortosa, Lupiáñez, et al., 2013; Tortosa, Strizhko, et al., 2013). Sin embargo, en lo referente al uso estratégico de la información personal para predecir el comportamiento de otras personas, ningún estudio se ha centrado en la influencia automática de las emociones cuando son irrelevantes y no predictivas.

Con el fin de profundizar en esto, se llevaron a cabo tres series experimentales adicionales. En la primera de ellas, nos centramos en la obtención a nivel comportamental de efectos de conflicto emocional en un contexto de interacción social. Aquí llevamos a cabo tres experimentos en los que se utilizó una versión modificada del Juego de la Confianza (*Trust Game*; Berg, Dickhaut, & McCabe, 1995) propuesta por Tortosa et al. (2013). Consiste en una tarea de tipo económico en la que los participantes deben decidir cuándo invertir su dinero basándose en la identidad del compañero como clave predictiva de su comportamiento en el juego y, en consecuencia, de los resultados más probables fruto de la interacción. Al comienzo de la tarea se instruye a los participantes en las tendencias cooperativas o no cooperativas de sus compañeros de juego. La ganancia económica de los participantes (real o ficticia, dependiendo del estudio), al final de cada ronda de juego, depende de la interacción, siendo mayor cuando el participante invierte su dinero y el compañero actúa de manera recíproca y menor cuando este último no coopera. En esta primera serie fueron manipulados el tipo de compañero, su expresión emocional (no predictiva), la validez de la información personal (porcentaje en el que predice realmente el comportamiento del compañero). En los dos primeros experimentos de la serie las instrucciones siempre acentuaron de manera explícita la necesidad de ignorar la emoción, dada su irrelevancia. En el tercero, sin embargo, no se ofreció instrucción alguna acerca de las emociones.

En la Serie III, aplicando una metodología similar a la utilizada en la Serie II, registramos la actividad cerebral de los participantes durante el desarrollo de la tarea empleando la técnica de resonancia magnética funcional (RMf). Esta serie nos permitió evaluar la implicación de las redes neurales relacionadas con el análisis del rostro (corteza extrastriada y giro fusiforme principalmente; ver Haxby & Gobbini, 2011; Haxby, Hoffman, & Gobbini, 2000), así como la respuesta neural de las estructuras involucradas en la detección e implementación de mecanismos de control (Botvinick et al., 2004; Egner et al., 2008; Egner & Hirsch, 2005) y su interacción con áreas relacionadas con el procesamiento de información social.

Por último, en la Serie IV evaluamos los niveles de procesamiento afectados por el conflicto emocional en contextos sociales mediante registros electrofisiológicos de alta densidad. Además, comparamos bloques en los que las emociones eran ignoradas (de forma similar a las dos series previas) con bloques en los que eran empleadas como los elementos predictivos clave del comportamiento de los compañeros. Esto nos permitió obtener un mapa temporal del procesamiento del conflicto emocional a lo largo de distintas etapas (indexadas por los potenciales P1, N1, N170, VPP, N2 y P3).

Las series experimentales han sido organizadas en capítulos diferentes. Las series I y II han sido publicadas en sendos artículos científicos. Las series III y IV están en fase de preparación para su posterior envío a revistas de impacto dentro del campo de estudio. Cada uno de los capítulos que siguen a continuación están estructurados como informes científicos, que incluyen las secciones de Resumen, Introducción, Métodos, Resultados, Discusión y referencias. Tras la presentación de los capítulos de cada serie, se concluirá con una discusión general de los principales resultados obtenidos, en relación al impacto que estos puedan tener en el estudio de los mecanismos de control.

2.3. Referencias

Barrett, L. F. (2012). Emotions are real. *Emotion (Washington, D.C.)*, *12*(3), 413-429. http://doi.org/10.1037/a0027555

Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a Psychological Primitive. *Advances in Experimental Social Psychology*, *41*, 167-218. http://doi.org/10.1016/S0065-2601(08)00404-8

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, *10*(1), 122-142. http://doi.org/10.1006/game.1995.1027

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. http://doi.org/10.1016/j.tics.2004.10.003

Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2008). Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(6), 1475-1484. http://doi.org/10.1093/cercor/bhm179

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, *8*(12), 1784-1790. http://doi.org/10.1038/nn1594

Eriksen, B. A., & Eriksen, C. W. (1974). Eriksen BA, Eriksen CW. Effects of noise letters upon identification of a target letter in a non-search task. Percept Psychophys 16: 143-149. *Perception & Psychophysics*, *16*(1), 143-149. http://doi.org/10.3758/BF03203267

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*(5), 349-358. http://doi.org/10.1037/0003-066X.43.5.349

Frijda, N. H., & Mesquita, B. (1994). The social roles and functions of emotions. En S. Kitayama & H. R. Markus (Eds.), *Emotion and culture: Empirical studies of mutual influence* (pp. 51-87). Washington, DC, US: American Psychological Association.

Garcia-Garcia, M., Domínguez-Borràs, J., SanMiguel, I., & Escera, C. (2008). Electrophysiological and behavioral evidence of gender differences in the modulation of distraction by the emotional context. *Biological Psychology*, *79*(3), 307-316. http://doi.org/10.1016/j.biopsycho.2008.07.006

Haxby, J., & Gobbini, M. I. (2011). Distributed Neural Systems for Face Perception. En A. Calder, G. Rhodes, M. Johnson, & J. Haxby (Eds.), *Oxford Handbook of Face Perception* (pp. 93–110). OUP Oxford.

Haxby, J., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223-233. http://doi.org/10.1016/S1364-6613(00)01482-0

Ibanez, A., Kotz, S. A., Barrett, L., Moll, J., & Ruz, M. (2014). Situated affective and social neuroscience. *Frontiers in Human Neuroscience*, *8*, 547. http://doi.org/10.3389/fnhum.2014.00547

Kanske, P., & Kotz, S. A. (2010a). Emotion Speeds up Conflict Resolution: A New Role for the Ventral Anterior Cingulate Cortex? *Cerebral Cortex*, bhq157. http://doi.org/10.1093/cercor/bhq157

Kanske, P., & Kotz, S. A. (2010b). Modulation of early conflict processing: N200 responses to emotional words in a flanker task. *Neuropsychologia*, *48*(12), 3661-3664. http://doi.org/10.1016/j.neuropsychologia.2010.07.021

Kanske, P., & Kotz, S. A. (2011). Positive emotion speeds up conflict processing: ERP responses in an auditory Simon task. *Biological Psychology*, *87*(1), 122-127. http://doi.org/10.1016/j.biopsycho.2011.02.018

Keltner, D., & Gross, J. J. (1999). Functional Accounts of Emotions. *Cognition and Emotion, 13*(5), 467-480. http://doi.org/10.1080/026999399379140

Moser, A., Gaertig, C., & Ruz, M. (2014). Social information and personal interests modulate neural activity during economic decision-making. *Frontiers in Human Neuroscience, 8,* 31. http://doi.org/10.3389/fnhum.2014.00031

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*(32), 11087-11092. http://doi.org/10.1073/pnas.0805664105

Parkinson, B. (1996). Emotions are social. *British Journal of Psychology (London, England: 1953), 87 (Pt 4),* 663-683.

Ruz, M., Ibanez, A., Kotz, S. A. E., Barrett, L., & Moll, J. (2014). *Interactions between emotions and social context - Basic, clinical and non-human evidence*. Frontiers E-books.

Ruz, M., Madrid, E., & Tudela, P. (2013). Interactions between perceived emotions and executive attention in an interpersonal game. *Social Cognitive and Affective Neuroscience*, *8*(7), 838-844. http://doi.org/10.1093/scan/nss080

Ruz, M., & Tudela, P. (2011). Emotional conflict in interpersonal interactions. *NeuroImage*, *54*(2), 1685-1691. http://doi.org/10.1016/j.neuroimage.2010.08.039

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science (New York, N.Y.)*, *308*(5728), 1623-1626. http://doi.org/10.1126/science.1110589

Tortosa, M. I., Lupiáñez, J., & Ruz, M. (2013). Race, emotion and trust: an ERP study. *Brain Research*, *1494*, 44-55. http://doi.org/10.1016/j.brainres.2012.11.037

Tortosa, M. I., Strizhko, T., Capizzi, M., & Ruz, M. (2013). Interpersonal effects of emotion in a multi-round Trust Game.

Whalen, P. J., Bush, G., Shin, L. M., & Rauch, S. L. (2006). The emotional counting Stroop: a task for assessing emotional interference during brain imaging. *Nature Protocols*, *1*(1), 293-296. http://doi.org/10.1038/nprot.2006.45

Capítulo 3: Serie Experimental I. Cognitive and affective control in a flanker word task: Common and dissociable brain mechanisms

El contenido de este capítulo ha sido publicado como Alguacil, S., Tudela, P. & Ruz, M. (2013). Cognitive and affective control in a flanker word task: common and dissociable brain mechanisms. *Neuropsychologia*, 51, 1663-1672.

Abstract

In the present study we compared the nature of cognitive and affective conflict modulations at different stages of information processing using electroencephalographic recordings. Participants performed a flanker task in which they had to focus on a central word target and indicate its semantic category (cognitive version) or its valence (affective version). Targets were flanked by congruent or incongruent words in both versions. Although tasks were equivalent at the behavioral level, event-related potentials (ERPs) showed common and dissociable cognitive and emotional conflict modulations. At early stages of information processing, both tasks generated parallel sequential conflict effects in the P1 and N170 potentials. Later, the N2 and the first part of the P3 wave were exclusively modulated by cognitive conflict, whereas the last section of the P3 deflection/LPC (Late Positive Component) was only involved in affective current conflict processing. Therefore, the whole data set suggests the existence of early common mechanisms that are equivalent for cognitive and affective materials and later task-specific conflict processing.

3.1. Introduction

When performing a task, humans often have to focus their attention on stimuli that are relevant to the task at hand and overcome the influence of distracting, irrelevant information. These situations require the implementation of control processes, given that relevant and irrelevant stimuli often trigger opposite action tendencies (see Norman & Shallice, 1986; Posner & Fan, 2004). How we exert control and solve such conflicts is a key research question that has been studied for many years using stimulus-response compatibility (SRC) tasks (Kornblum, Hasbroucq & Osman, 1990). In this type of settings, goal-relevant targets are surrounded by irrelevant stimuli that are associated with the same (congruent) or opposite (incongruent) action tendencies. The typical conflict effect (Eriksen & Eriksen, 1974) arises, as responses are faster and more accurate in congruent than in incongruent trials. On the other hand, the conflict adaptation effect (Gratton, Coles & Donchin, 1992) is evaluated based on the interaction between the previous and current congruency in consecutive trials.

Responses on trials following an incongruent trial show decreased conflict. This finding is supposed to reflect the heightened activation of control mechanisms, which seem to be more available for the resolution of new incongruences (Botvinick, Braver, Barch, Carter & Cohen, 2001; Egner, Etkin, Gale & Hirsch, 2008). This effect seems to be domain-specific, as it does not transfer across different SRC tasks that are performed sequentially (Egner, 2008; Egner et al., 2008; Funes, Lupiáñez & Humphreys, 2010).

In neural terms, according to the conflict-monitoring hypothesis, the anterior cingulate (ACC) and dorsolateral prefrontal (dlPFC) cortices are involved in monitoring conflictive situations and implementing control mechanisms, respectively, when required (Botvinick, Nystrom, Fissell, Carter & Cohen, 1999; Cohen, Botvinick & Carter, 2000; MacDonald, Cohen, Stenger & Carter, 2000). This hypothesis posits that the ACC monitors the occurrence of conflict online and the dlPFC holds task-relevant representations and amplifies cortical responses to reduce conflict. After that, the ACC is thought to provide feedback (Botvinick et al., 1999; Botvinick et al., 2001; Cohen et al., 2000; Egner & Hirsch, 2005).

Initially, research on emotional conflict used SRC paradigms such as the emotional counting Stroop task (Whalen et al., 1998) and other related tasks (Bush, Luu & Posner, 2000) in which emotional information was a distractor but not a target, as targets were always cognitive. However, further developments were aimed at equating cognitive and affective types of conflict within a single task (Egner et al., 2008; Kanske & Kotz, 2010a) and using the same type of stimuli across conditions (Ochsner, Hughes, Robertson, Cooper & Gabrieli, 2008).

For example, Egner et al. (2008) modified the classic Stroop task (Stroop, 1935) by displaying overlapping male and female faces (expressing fear or happiness) and words (referring to the gender or the emotional expression of the faces). Participants had to respond to the gender of the faces in the cognitive version of the task and to the emotional expression of the faces in the affective version of the task. In both cases, participants had to ignore irrelevant superimposed verbal information. Thus, faces and words could be either congruent or incongruent in both cognitive and emotional tasks. Similarly, Ochsner et al. (2008) used a task in which a word (target) was presented at the center of the screen, flanked by another two words displayed above and below. Participants had to indicate the semantic category (fruits vs. metals) of the target in the cognitive version of the task and respond to the valence of the words (positive vs. negative) in the emotional version.

To date, several reports have jointly explored the neural substrate of cognitive and affective conflict. Results converge on the finding of common areas of activation for both types of tasks in the dorsal ACC. However, differences have also been found between cognitive and affective versions.

Cognitive conflict seems to involve activity in the dorsolateral prefrontal cortex, whereas affective conflict seems to recruit the rostral medial prefrontal and ventral anterior cingulate cortices (Bush et al., 2000; Egner et al., 2008; Kanske & Kotz, 2010a; Ochsner et al., 2008). These results are mainly derived from functional magnetic resonance imaging (fMRI). This technique, however, is not appropriate to conduct analyses on how different types of conflict affect the various rapid stages of information processing that take place from stimulus onset to response execution. By contrast, electroencephalography (EEG) provides excellent temporal resolution that helps understand whether cognitive and emotional conflicts generate common or dissociable effects along the various stages of stimulus processing.

Several event-related potentials (ERPs) are sensitive to conflict modulation across different stages of information processing. First, the P1 and N170 are early potentials that reflect perceptual processing in visual cortices (Hillyard, 2009; Luck, Woodman & Vogel, 2000). The amplitude of the P1 is also heightened by selective attention to spatial positions and stimulus features (Hillyard, 2009; Luck, 2005), and seems to reflect the amplification of relevant information (Hillyard, Vogel & Luck, 1998).

According to research conducted by Meeren, Heijnsbergen and Gelder (2005), the P1 displays larger amplitudes for congruent than for incongruent face-body emotional expressions. Similarly, Scerif, Worden, Davidson, Seiger and Casey (2006) used a classic flanker paradigm and found conflict adaptation effects in the P1 potential. At that stage, incongruent trials preceded by trials of the same type had larger amplitude than the other combinations of previous and current congruency in a sequence of trials.

Together, these results suggest that conflict and its previous congruency context can influence the processing of perceptual information at a very early stage. The N170 potential is associated with perceptual discrimination and object categorization processes (Hillyard, 2009; Luck et al., 2000) and can be modulated by attention (see Ruz & Nobre, 2008a; Aranda, Madrid, Tudela & Ruz, 2010). In two recent parallel experiments, Zhu, Zhang, Wu, Luo and Luo (2010) used a Stroop face-word manipulation in which participants had to indicate the emotional category of the target stimulus (the face in Experiment 1 and the word in the Experiment 2) while ignoring the distractor element (similar to Egner et al., 2008). In Experiment 1 (in which participants responded to the face), the N170 potential had larger amplitude in the incongruent condition than in the congruent condition. In Experiment 2 (in which the word was the target), the amplitude pattern was reversed. Thus, the amplitude of this potential was enhanced by task-relevant information during emotional conflict processing (Zhu et al., 2010). Altogether, this evidence from the P1 and N170 potentials shows that cognitive and emotional conflict can modulate the perceptual processes reflected in these potentials and – at least in the P1 potential – reflects the control of cognitive conflict according to its previous congruency context.

The term 'N2' includes several waves that have generally been related to the allocation of attention to target stimuli and the suppression of irrelevant stimuli (Luck, 2005). The N2 that is most strongly associated with cognitive control is a negative deflection that takes place around 200 ms after stimulus onset with a fronto-central topography (Heil, Osman, Wiegelmann, Rolke & Hennighausen, 2000; Kopp, Rist, & Mattler, 1996). Using a classic flanker task, Kopp et al. (1996) found that the amplitude of the N2 was larger in incongruent trials than in congruent trials (Kopp et al., 1996; Van Veen & Carter, 2002). This result, which seems to reflect the implementation of control mechanisms (Folstein & Van Petten, 2008), has since been replicated several times (Folstein & Van Petten, 2008; Van Veen & Carter, 2002). With the help of computer simulations, Yeung et al. (Yeung, Botvinick & Cohen, 2004) proposed that the N2 potential reflects conflict monitoring in correct responses prior to response execution. In fact, the N2 potential seems to reflect a process of detection of potential future errors that takes place before response execution (Yeung et al., 2004). Regarding emotions, Kanske and Kotz (2010b; 2011a,b) have used various paradigms to study the influence of emotional and neutral materials during the resolution of cognitive conflict. In their tasks (Kanske & Kotz 2010b; 2011a,b), however, the valence of the stimuli is not relevant for performance, and therefore the response never concerns this dimension. In two of their studies, for example, Kanske and Kotz (2010b; 2011b) used a flanker task composed of affective and neutral words. In each trial, three stimuli (a target word flanked by two additional words) were displayed in the same (congruent condition) or different colors (incongruent condition). Participants were required to indicate the color of the target but the meaning of the words was irrelevant. The presence of irrelevant affective material generated an N2 of larger amplitude compared to irrelevant words with neutral meaning. This effect was accompanied by a decrease of the conflict effect for affective materials. The authors therefore concluded that the emotional nature of irrelevant stimuli speeds up conflict resolution.

Finally, the P3 potential has been associated with stimulus selection and categorization and with the implementation of the required response (Luck, 2005). Polich (2006) posited that the P3 would be modulated by inhibitory processes involved in avoiding ‘extraneous’ information and in focusing attentional resources on the relevant elements of the task. Thus, the harder it was to inhibit ‘extraneous’ elements, the lower the amplitude of the P3. In addition, studies using SRC paradigms have reported larger amplitudes of this potential for congruent than for incongruent conditions (Neuhaus et al., 2010; Valle-Inclán, 1996a). Although few experiments have explored the susceptibility of the P3 to emotional conflict, a study using words in a modified Go/NoGo task found larger P3 amplitudes in NoGo trials but only for affective and nonneutral words (Chiu, Holmes & Pizzagalli, 2008).

In summary, previous studies comparing cognitive and emotional conflict resolution have mainly used fMRI methods, and those recording EEG have not compared both types of material using the same type of stimuli and equating task demands. Therefore, in study we explored the temporal dynamics of cognitive and emotional conflict resolution using the same kind of materials and task settings across conditions. To this end, we adapted the paradigm developed by Ochsner et al. (2008) and recorded EEG while participants assessed the semantic category or affective valence of words flanked by congruent or incongruent stimuli in different blocks. We focused on the P1, N170, N2 and P3 potentials as markers of different stages of information processing and studied how current and previous trial congruency factors affected such stages. Based on previous fMRI studies, we predicted that both tasks would show both similarities and differences in the neural stages modulated by cognitive and emotional conflict. More specifically, we expected conflict to be reflected in larger P1 and N170 amplitudes for incongruent than for congruent trials. In addition, we expected to find a congruency effect in the N2 potential, with more negative amplitude for incongruent than for congruent trials. Finally, we hypothesized a conflict effect for both types of tasks at later P3 stages.

3.2. Methods

3.2.1. Participants

The sample was composed of 26 healthy volunteers (1 left-handed participant, 11 men, mean age: 22 years) with normal or corrected-to-normal vision and Spanish as their native language. Participants were recruited from the University of Granada in exchange for course credit. All participants signed a consent form approved by the local Ethics Committee.

3.2.2. Stimuli, apparatus and procedure

Participants performed a modified version of the classic flanker task (Eriksen & Eriksen, 1974), previously adapted by Ochsner et al. (2008). The cognitive and affective versions of this task were presented in different blocks. A group of 20 words was selected from the Spanish version of the ANEW (Affective Norms for English Words; Redondo, Fraga, Padrón & Comesaña, 2007) to serve as stimuli; these were divided in two 10-words sets. The cognitive task included 5 words of each semantic category (5 animals, 5 pieces of furniture and fittings) and the affective task included 5 words of high positive valence and 5 words of high negative valence. Their combination yielded 72[1] target-flanker sets, 36 for each task. Every possible target-flanker combination appeared only once in each block to avoid immediate stimulus repetitions. No further constraints were applied on stimuli, response repetitions or trial sequences. Words were matched across tasks in mean frequency of use (MnFreq = 14; SD = 2.13), number of letters (MnNLett= 6.4; SD=0.16) and mean concreteness (MnCon = 5.82; SD = 0.53), all *ps* > .05). The average valence was equal for words used in the cognitive (MnValCog = 5.42; SD = 1.09) and affective tasks (MnValAffe = 4.76; SD = 2.97) tasks, $F < 1$. Affective words had extreme negative (MnValNeg = 1.96 (DT = 1.49) and positive (MnValPos = 7.55; SD = 1.52) values; $F(1,8) = 368.88$, $p <. 001$.

The task was created and displayed using E-Prime 2.0 Professional software (Schneider, Eschman & Zuccolotto, 2002). All stimuli were composed of white uppercase letters displayed on a black background. Each trial comprised the following sequence (see Figure 1): after a fixation point (0.43°) of 2500 ms duration on average (random 2000-3000 ms range), target

[1] Six combinations from each set were removed to avoid the simultaneous presence of the same word as target and flanker.

and flanker words (2.57º / 3.42º on average) were displayed vertically aligned on the center of the screen for 1500 ms. From target onset, participants had 1500 ms to respond by pressing one of two buttons with their left and right index fingers. The association between hand and response was counterbalanced across tasks and participants.

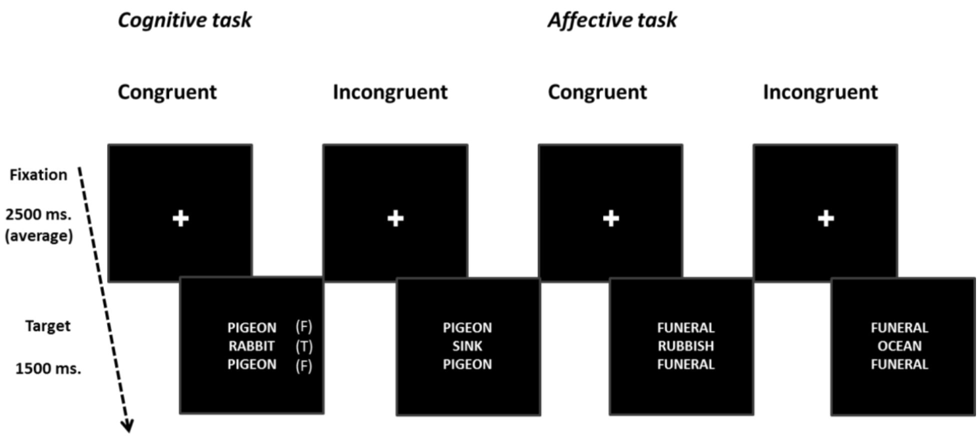

Figure 1: Experimental conditions and sequence of events in a trial. F = Flanker; T = Target. Target and flanker words have been translated into English.

Participants were asked to respond to the semantic category of the words in the cognitive version (animals vs. furniture) and to their valence in the affective task (positive vs. negative). In every trial, three words were displayed in the center of the screen (see Fig. 1). The central word was the target and was surrounded above and below by two other words (flankers). Participants were asked to focus on the target word and to ignore the flankers. Trials were congruent if the semantic category (cognitive task) or the valence indexes (affective task) were the same for target and flankers and incongruent if they were different.

At the beginning of the session, participants performed a training block of 30 trials (15 for each task) to familiarize themselves with the procedure. During this practice, they received feedback of their performance. The experiment was composed of 10 blocks (5 for each task) of 72 trials each (720 trials in total; 360 trials per task, each composed of 180 congruent and 180 incongruent trials), separated by short breaks. The order of presentation of each group of blocks (cognitive or affective) was counterbalanced across participants.

3.2.3. EEG recording and data analyses

Subjects were seated in front of the computer monitor in an electrically shielded room and were instructed to avoid eye blinks and movements during stimulus presentation and responses. EEG was recorded with a high-density 128-channel EEG system Geodesic Sensor Net (Tucker, 1994), referenced to the vertex channel. The head coverage included sensors lateral to and below both eyes to monitor horizontal and vertical eye movements (HEOG and VEOG). The EEG net was connected to an AC-coupled, high-input impedance amplifier (200 MΩ). At the beginning of the recording session, impedance was measured for each channel and was set below 50 kΩ, as recommended for Electrical Geodesics high-input impedance amplifiers. Amplified analog voltages (0.1 to 100-Hz band pass) were digitized at 250 Hz (12 bits A/D converter and 0.02 μV minimum resolvable voltage).

The continuous EEG was filtered offline using a 40 Hz low-pass filter. After that, the EEG was segmented 200 ms before and 800 ms after target onset. Subsequently, segments were submitted to software processing for identification of artifacts. Trials containing eye blinks or eye movements (electro-oculogram channel differences greater than 70 μV) or more than 20% of bad channels were not included in the ERPs. Data from consistent bad channels were later replaced using a spherical interpolation algorithm (Perrin, Pernier, Bertrand & Echallier, 1989). ERPs were re-referenced offline to the average. The 200 ms pre-stimulus epoch served as the baseline. A minimum criterion of 30 artifact-free trials per subject and condition (cognitive and affective CC, CI, IC, and II) was established to maintain an acceptable signal-to-noise ratio. A final grand-average was obtained per condition by pooling the subject's averages in each experimental condition. Eight group-average ERP waveforms were constructed according to task (*cognitive* vs. *affective*), current congruency (*congruent* vs. *incongruent*) and previous congruency (*congruent* vs. *incongruent*) factors. Conditions were equated in number of trials, $F<1$.

To facilitate the selection of spatio-temporal windows for amplitude analyses, the ERPs were submitted to a segmentation analysis performed with Cartool software (D. Brunet, Geneva, Switzerland; http://sites.google.com/site/fbmlab/cartool). This analysis shows the voltage distribution (topographies) over the scalp for each experimental condition as a function of time. These topographic maps represent stable periods of electrical field patterns that correspond to dissociable functional states of the brain (Murray, Brunet & Michel, 2008; Pascual-Marqui, Michel & Lehmann, 1995; Ruz & Nobre, 2008b), which normally match identifiable ERPs on the grand-averaged waveforms (Pascual-Marqui et al., 1995).

The topographic analysis was conducted with some limitations: the maps had to remain stable for 20 ms and the correlation between different maps could not be higher than 92% (see Aranda et al., 2010; Baines, Ruz, Rao, Denison & Nobre, 2011; Cojan, Archimi, Cheseaux, Waber & Vuilleumier, in press; Correa, Rao & Nobre, 2009; Foxe, Murray & Javitt, 2005; Morand, Harvey & Grosbras, in press; Nobre, Griffin & Rao, 2008; Ruz & Nobre, 2008b, for similar constraints). This procedure was used to obtain a common sequence of topographies, which served to select the spatio-temporal windows in the ERPs for data analyses.

The optimal sequence of topographies and number of maps (restricted by the constraints mentioned above) that best explained the entire data set were defined by a researcher-independent cross-validation (CV) criterion (Pascual-Marqui et al., 1995). This criterion is used to assess the validity of regression models (Snee, 1977). The topographic solution with the minimum CV value is chosen, as it is the one in which the smallest number of maps explains the greatest variance in the data. After that, the temporal ranges of the cluster maps or topographies are used to delimit the temporal windows for the ERP analyses.

Voltage analyses were performed on the spatio-temporal windows that captured the grand-average peaks of the potentials of interest. The selected electrodes were those where the components were maximally distributed (see Figures 4 and 5). The temporal windows were centered on the peak of the potentials in the grand-averaged waveforms and always fell within the time intervals of the corresponding topographic maps. Only correct responses were considered; errors, post-error trials and the first trials of each block were removed from the set. Data from three participants were rejected because they had too many bad channels or excessive artifacts during recording (they were also excluded from the behavioral analyses). Mean amplitude voltages averaged over the selected channels and time windows were submitted to repeated-measures ANOVAs with the same factors as those included in the group-average ERP. Analyses were carried out for the P1, N170, N2 and P3 potentials ($P3_1$, $P3_2$ and $P3_3$ included topographic constraints). The hemisphere was included as a factor (right, left) for the P1 and the N170. The same ANOVA was used with participants' mean errors and RTs.

3.3. Results

3.3.1. Behavioral results

Only correct responses were considered in the RT and the ERP analyses. Errors and post-error trials (6.42 %) were eliminated from the data. RTs shorter than 300 ms (0.82%) were excluded. In addition, trials at the beginning of each block (1.38 %) were also discarded.

Overall accuracy was 96%. Performance was more accurate in the congruent (Mn_{AccCon} = 97.9 %; SD = 1.5) than in the incongruent condition (Mn_{AccIn} = 96.9 %; SD = 1.8), $F(1,25) = 22.681$, $p < .001$. The interaction between current and previous congruency only reached marginal significance levels, $F(1,25) = 3.411$, $p = .077$[2]. All other comparisons were not significant, *ps* > 0.05.

---

[2] Subsequent comparisons did not show significant effects of the previous congruency factor for either congruent or incongruent current trials, all *ps*>0.26.

RT analyses showed a main effect of congruency, $F(1,25) = 33.627$, $p<.001$, as responses were faster in the congruent ($Mn_{RTCon} = 639.05$; SD = 96.60) than in the incongruent condition ($Mn_{RTInc} = 653.84$; SD = 102.98); see Figure 2. The previous congruency factor was close to significance levels, $F(1,25) = 3.343$, $p=.079$ ($Mn_{RTPrevCon} = 644.87$; SD = 98.11 vs. $Mn_{RTPrevInc} = 647.97$; SD = 101.39). There were no other significant effects, all $Fs < 1$. The mean and standard deviation of behavioral data in all conditions are shown in Table 1.

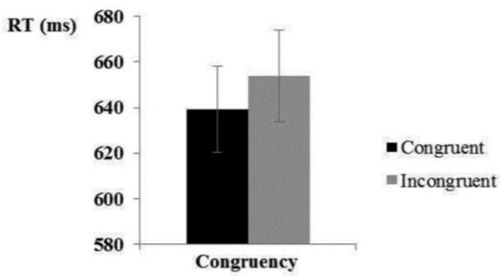

Figure 2: Congruency effects on RTs.

Table 1: Descriptive statistics of behavioral data

	COGNITIVE TASK		AFFECTIVE TASK	
PREVIOUS CONGRUENT	**Mean ACC (SD)***	**Mean RT (SD)**	**Mean ACC (SD)**	**Mean RT (SD)**
Current Congruent (**CC**)	0.98 (0.14)	638.92 (177.22)	0.98 (0.13)	636.12 (170.89)
Current Incongruent (**CI**)	0.97 (0.18)	651.55 (187.57)	0.97 (0.17)	652.87 (179.85)
PREVIOUS INCONGRUENT	**Mean ACC (SD)**	**Mean RT (SD)**	**Mean ACC (SD)**	**Mean RT (SD)**
Current Congruent (**IC**)	0.98 (0.15)	643.71 (184.43)	0.98 (0.15)	637.51 (172.99)
Current Incongruent (**II**)	0.97 (0.17)	658.02 (187.63)	0.97 (0.16)	652.98 (177.62)

* SD= Standard Deviation

3.3.2. Electrophysiological results

The topographic analysis showed that ERP data was best explained by 7 sequential maps, which are displayed in Figure 3.

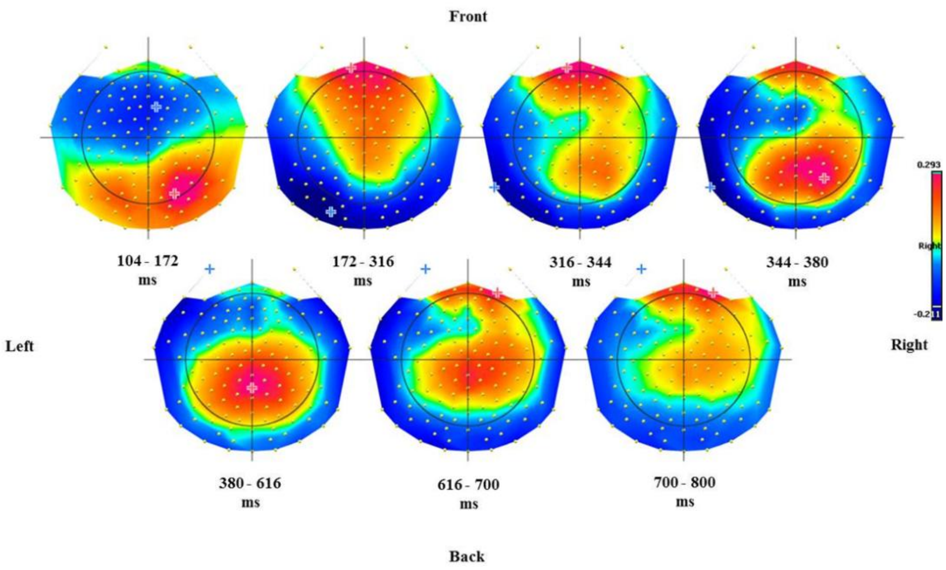

Figure 3: Sequence of topographies associated with the task.

The first map showed negative polarity in frontal regions and positive polarity in posterior regions, corresponding to the P1 potential. The second map showed positive voltages at anterior-central sites and negative voltages at posterior sites, and corresponded to the N170. The fourth map had negative values in frontal regions and positive voltages in central ones, corresponding to the N2 potential. The seventh, eighth and ninth maps showed positive polarity at central sites. These last three maps corresponded to first ($P3_1$), second ($P3_2$) and third sections ($P3_3$) of the P3 potential, respectively. Each topography is thought to represent a different processing stage and dissimilarities in the underlying brain sources (Lehmann, 1987). Hence, each segment of the P3 wave seems to correspond to a different brain microstate (Pascual-Marqui et al., 1995; Michel, Henggeler & Lehmann 1992). As a result, the segments were analyzed separately. Similar strategies and map labeling can be found in Baines et al. (2011) and in Nobre et al. (2008).

3.3.2.1. Early processing stages

The P1 potential peaked approximately at 138 ms and was analyzed from 106 ms to 170 ms over posterior bilateral electrodes (see Figure 4). The analysis of the averaged amplitudes showed a significant hemisphere effect, $F(1,25) = 13.013$, $p < .05$, with larger positive voltages in right electrodes ($Mn_{AmpRH} = 1.90$; SD = 2.03) than in left electrodes ($Mn_{AmpLH} = 0.89$; SD = 1.52). In addition, the interaction between current and previous congruency factors was significant, $F(1,25) = 5.377$, $p < .05$. Subsequent comparisons showed an effect of previous congruency only for current incongruent trials, $F(1,25) = 4.228$, $p = .05$. In other words, incongruent trials following congruent trials showed larger positive amplitudes ($Mn_{AmpCI} = 1.48$; SD =1.67) than incongruent trials preceded by incongruent trials ($Mn_{AmpII} = 1.29$; SD =1.60). This difference was not significant for current congruent trials, *p*

= 0.347 (see Figure 4). The cognitive vs. affective nature of the task did not show an effect or interact with any other factor, all *Fs* < 1.

The N170 potential peaked approximately at 215 ms and was analyzed from 175 ms to 255 ms over posterior bilateral electrodes (see Figure 4). The analysis of the averaged amplitude for the N170 showed an effect of hemisphere, $F(1,25) = 26.379$, $p<.001$, as more negative amplitudes were found in left ($Mn_{AmpLH} = -2.07$; SD = 1.80) than right locations ($Mn_{AmpRH} = -0.83$; SD =1.79). In addition, the interaction between current and previous congruency factors was again significant, $F(1,25) = 13.792$, $p =.001$. Planned comparisons showed an effect of the previous congruency factor only for current incongruent trials, $F(1,25) = 9.402$, $p<.01$; the voltage values were more negative for incongruent trials following identical trials ($Mn_{AmpII} = -1.56$; SD = 1.65) than for those following congruent ones ($Mn_{AmpCI} = -1.40$; SD = 1.67). This difference was not significant for current congruent trials, $F < 1$; see Figure 4. Again, neither the cognitive vs. affective nature of the task nor its interactions with other factors were significant, all $ps > .05$.

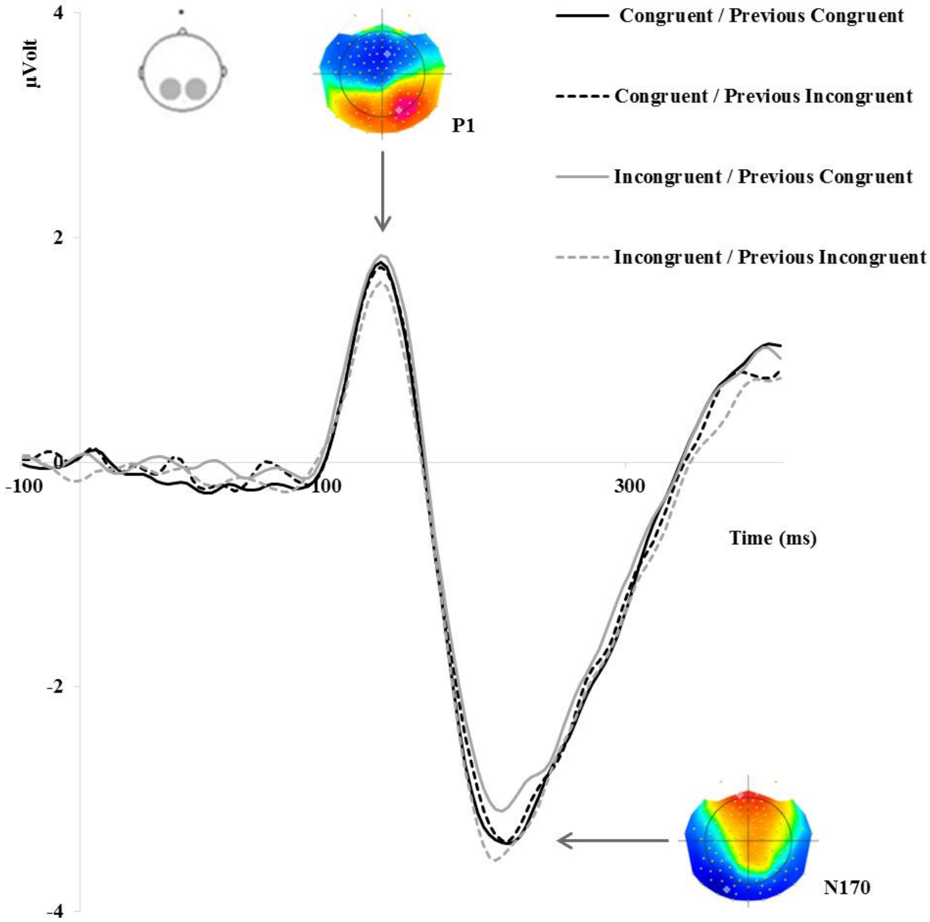

Fig. 4: Topographies during the P1 (top) and N170 time windows (bottom), location of the selected channel (next to T5 and P3) and the ERP of a representative posterior electrode. Positivity is plotted upwards in all figures.

3.3.2.2. Intermediate processing stages

The N2 potential peaked approximately at 330 ms and was analyzed from 318 ms to 342 ms over central electrodes. The ANOVA with the averaged amplitude values showed an effect of previous congruency, $F(1,25) = 8.808$, $p < .01$, as trials following an incongruent trial showed a more negative-going deflection ($Mn_{AmpPrevInc}$ = 0.52; SD = 1.67) than those preceded by a congruent trial ($Mn_{AmpPrevCon}$ = 0.79; SD = 1.54). Two significant interactions were also found: task x current congruency, $F(1,25) = 5.606$, $p < .05$, and task x previous congruency, $F(1,25) = 5.970$, $p < .05$. Subsequent analyses showed that there was an effect of current congruency in the cognitive task, $F(1,25) = 6.151$, $p < .05$, as the N2 amplitude in the incongruent condition was more negativegoing (Mn_{AmpInc} = 0.60; SD = 1.80) than in the congruent condition (Mn_{AmpCon} = 0.92; SD = 1.70); see Figure 5. Moreover, there was an effect of previous congruency again only in the cognitive task, $F(1,25) = 21.103$, $p < .001$; this was because trials following an incongruent trial showed a more negative-going deflection ($Mn_{AmpPrevInc}$ = 0.47; SD = 1.80) than those following a congruent trial ($Mn_{AmpPrevCon}$ = 1.04; SD = 1.69). There were no effects of current or previous congruency in the affective task, all *Fs < 1*.

3.3.2.3. Late processing stages

Following the topographic maps associated with the P3 wave (see Figure 5), this potential was analyzed in three separate time windows ($P3_1$, from 382-614 ms; $P3_2$, from 616-700 ms; and $P3_3$, from 700-800 ms) over the same central channels. The $P3_1$ reflected an effect of previous congruency, $F(1,25) = 12.054$, $p < .01$, as the amplitude of trials following congruent trials was larger ($Mn_{AmpPrevCon}$ = 2.53; SD = 2.11) than that of those following incongruent trials ($Mn_{AmpPrevInc}$ = 2.25; SD = 2.13). There was also a marginal interaction between task and current congruency, $F(1,25) = 4.075$, $p = .05$. A current congruency effect was found only in the cognitive task, $F(1,25) = 4.515$, $p < .05$, as the amplitude of $P3_1$ was larger in the congruent condition (Mn_{AmpCon} = 2.49; SD = 2.20) than in the incongruent one (Mn_{AmpInc}=2.25; SD =2.18). This effect was absent in the affective task, $F < 1$ (see Figure 5).

The analysis of the second time window ($P3_2$) yielded a main effect of previous congruency, $F(1,25) = 6.228$, $p < .05$, as trials following a congruent trial showed a more positive amplitude ($Mn_{AmpPrevCon}$ = 2.32; SD = 1.98) than those following an incongruent trial ($Mn_{AmpPrevInc}$ = 2.10; SD = 2.02). There were no other significant effects; neither the task factor nor its possible interactions reached significant values, all *ps* > .05. The ANOVA of the averaged amplitudes in the $P3_3$ showed, as in the $P3_1$ analysis, an interaction between task and current congruency, $F(1,25) = 4.724$, $p < .01$. The analysis of the affective task showed an effect of current congruency, $F(1,25) = 8.469$, $p < .01$. This was because the $P3_3$ had a more positive amplitude in the incongruent (Mn_{AmpInc} = 1.39; SD = 1.91) than in the congruent condition (Mn_{AmpCon} = 1.06; SD = 1.74). The effect of current congruency was absent in the cognitive task, $F < 1$. There was an effect of previous congruency, *F*(1,25) = 5.711, *p* < .05, which was better explained by its interaction with the task factor, *F*(1, 25) = 4.350, *p* < .05. Planned comparisons for each task showed an effect of previous congruency only for the cognitive task, *F*(1,25) = 10.101, *p* < .01, which showed larger amplitudes in trials preceded by a congruent one ($Mn_{AmpPrevCon}$ = 1.33; SD = 1.62) than in those preceded by an incongruent trial ($Mn_{AmpPrevInc}$ = 0.93; SD = 1.72). The effect of previous congruency was absent in the affective task, *p* > .05. There were no other effects of interest, all *ps* > .05; see Figure 5.

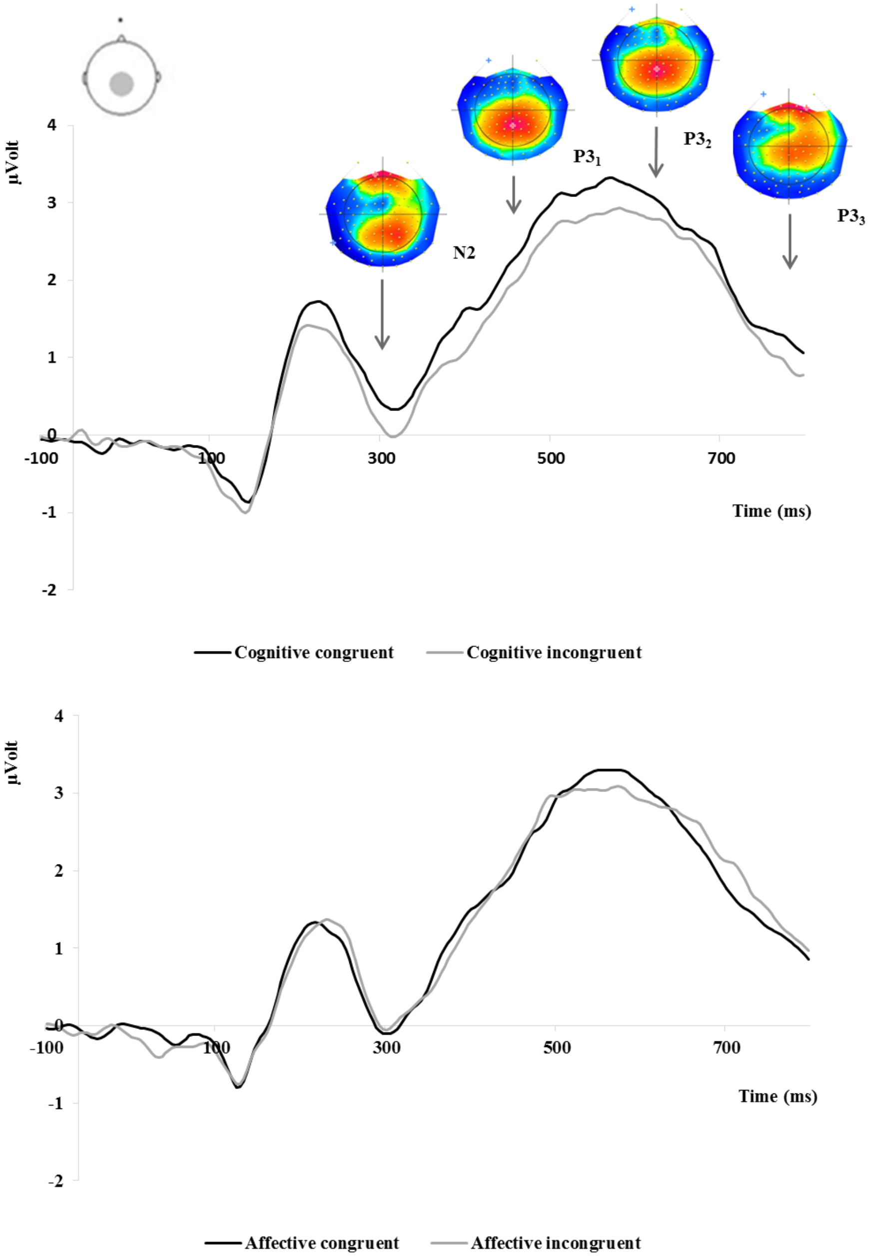

Fig. 5: The display shows the topography (top panel) of the experimental tasks during the N2 (left) and P3 ($P3_1$, $P3_2$ and $P3_3$) time windows (upper panel: cognitive task; bottom panel: affective task), the location of selected channels and the ERP of a representative central electrode (below Cz) for the current congruency effect in these potentials.

3.4. Discussion

The aim of the present study was to extend our knowledge of the differences and similarities in the modulations generated by cognitive and emotional conflict processing and their contextual effects in stimulus processing stages. Our behavioral results showed a higher rate of errors and slower responses in incongruent trials. This slowing down of responses seemed to persist in consecutive trials, as indicated by an effect of previous congruency, although this effect did not reach significance levels in the behavioral data. The set of results of previous and current congruency factors seemed to be equivalent in cognitive and affective tasks. In addition, electrophysiological data suggest the existence of parallel conflict modulations for both kinds of stimuli at initial stages of processing and later dissociations according to the cognitive or affective nature of the task.

Our behavioral results are consistent with previous studies using SRC paradigms with cognitive and affective materials. We replicated the typical conflict effect (Eriksen & Eriksen, 1974; Botvinick et al., 2001). However, we did not observe conflict adaptation effects in the interaction between current and previous congruency. The reasons for this lack of replication at the behavioral level are unclear. One potential explanation may lie in the differences between the current paradigm and previous SCR tasks. In classic tasks, participants are asked to judge simple features of stimuli, such as their spatial orientation. In the current task, however, participants had to evaluate the meaning of the words. Thus, the heightened cognitive control recruited by incongruent conditions may have not facilitated further conflict resolution but rather set the system in a ‘more careful’ mode, which would slow down subsequent responses regardless of trial type. This is reminiscent of some observations in the literature about the slowing down of responses owing to control mechanisms: the *post-error slowing effect* and the *bivalence effect,* also known as *post-conflict slowing effect*. The post-error slowing effect happens when participant responses slow down after an error (Botvinick et al., 2001). In the *bivalence effect,* responses to univalent stimuli slow down if they are presented in blocks also containing incongruent bivalent stimuli (Grundy, Benarroch, Woodward, Metzak, Whitman & Shedden, 2011; Woodward, Meier, Tipper & Graf, 2003; Woodward, Metzak, Meier & Holroyd, 2008) or if they occur immediately after an incongruent bivalent trial (Verguts, Notebaert, Kunde & Wühr, 2011). Thus, as Verguts et al. (2011) noted, task difficulties or interruptions such as incongruencies seem to lead to slower task performance in both current congruent and incongruent trials, as is the case in our results.

The P1 and N170 potentials showed conflict adaptation effects for both tasks, although these effects were absent in behavioral indices. Conflict adaptation effects had been previously reported for the P1 (Scerif et al., 2006), in an experiment in which arrow targets were preceded by three consecutive congruent or incongruent trials. Scerif et al. (2006) found that incompatible trials preceded by an incompatible context generated a P1 of larger amplitude than the rest of possible combinations of previous context and current conflict. In their explanation, the sequence in the incongruent context helped attention focus on the location of the target, thus reducing flanker interference in the following incongruent trial (Scerif et al., 2006). We found conflict adaptation effects at early processing stages too, which mimicked the results obtained by Scerif et al. (2006) in the N170 potential but showed the reverse pattern for the P1. In both cases, conflict only modulated the amplitude of the visual potentials in current incongruent trials.

However, whereas the amplitude of the P1 was larger for incongruent trials preceded by congruent ones, the N170 was largest for a sequence of two incongruent trials. In contrast to the arrows used by Scerif et al. (2006), participants in our study had to judge the semantic nature of the target words, which calls for a more elaborated processing. While responding to the spatial orientation of arrows depends heavily on the allocation of attention and the spatial arrangement of the flanker and targets, these elements seemed less important for the correct performance of our task or may be more relevant at later stages of processing (e.g. Ruz & Nobre, 2008a).

If heightened amplitude is taken as an index of deeper or more attentive processing (Hillyard et al., 1998; Luck, Fan & Hillyard, 1993), our results suggest the following: initially, the perception of incongruent targets is boosted when they follow a congruent context (P1 potential) but a few milliseconds later incompatible trials following identical ones are subject to heightened processing (N170 potential); this facilitates the discriminative processing of the stimuli (Hillyard, 2009; Luck et al., 2000). In conflict terms, these sequential patterns could be interpreted as the detection of conflict (P1) and the subsequent implementation of control (N170), respectively. According to the conflict monitoring theory (Botvinick et al., 1999; Egner, 2007; Gratton et al., 1992), current incongruent trials preceded by congruent ones are supposed to reflect a situation low in control (as the previous trial was congruent) but high in conflict (as the current situation is conflictive). The fact that the P1 potential had the largest amplitude in this situation for the cognitive and emotional versions of the task suggests that in both cases the P1 reflected the detection or monitoring of conflict. By contrast, current incongruent trials preceded by others of the same type are thought to reflect a situation of high control and low conflict. As the amplitude of the N170 potential was heightened in this trial sequence, it could be argued that this modulation reflected the implementation of control processes for cognitive and emotional conflicts alike (Botvinick et al., 1999; Egner, 2007).

The lack of sequential effects at the behavioral level argues for caution in the interpretation of the ERP results. It should be noted that Ochsner et al. (2008) used the same paradigm but did not show conflict adaptation (either in behavioral or in neuroimaging data). This suggests that the present paradigm is not best suited to obtain behavioral adaptation effects. However, the electrophysiological findings strongly suggest that these processes take place for both types of conflict at *early* processing stages, which makes them even more intriguing. Electrophysiological data also indicate that response execution stages were no longer affected by adaptation effects (as there were no interactions between current and previous congruency for either task in the P3), which could help explain the lack of adaptation effects in behavioral indexes. In any case, the P1 and N170 results cannot be explained by spurious stimuli repetitions affecting congruent and incongruent trials in a differential manner, as the two conditions were equivalent at the perceptual level and every possible target-flanker combination in our experiment appeared only once in each block. Moreover, as mentioned by Egner (2007), the repetition account is less plausible when a large set of stimuli is used, as is the case in our experiment.

Another relevant outcome from the P1 and N170 potentials was the absence of main or interactive effects of the cognitive vs. affective nature of the task. Although the words used in the two conditions were different, they were equated in average levels of item length, frequency, concreteness and valence, which could help explain the lack of main task effects. Moreover, as suggested by Olofsson, Nordin, Sequeira and Polich (2008), the processing of affective materials at early stages seems to be more sensitive to negative or unpleasant items when they are confronted with positive or neutral ones. In our manipulation, however, these materials only appeared separately in congruent trials but not in the incongruent ones; hence, we cannot evaluate the conflict effects independently for each valence category. Nevertheless, the outcome of equivalent conflict adaptation effects across the cognitive and affective versions strongly suggests that these early mechanisms are general and common to different task requirements.

The N2 potential, which has been associated with cognitive control using several kinds of paradigms (for a review, see Folstein & Van Petten, 2008), displayed larger negative amplitudes for incongruent than congruent trials during the cognitive task. This pattern is consistent with previous evidence of N2 modulations by focusing attention on targets and filtering out irrelevant stimuli (Luck, 2005). More specifically, an extensive body of literature has reported similar results, using mostly cognitive materials

(Folstein & Van Petten, 2008; Kanske & Kotz, 2010b, 2011a; Kopp et al., 1996). N2 modulations seem to be related to the response. Therefore, they are thought to reflect the monitoring of inappropriate planned responses during conflictive situations mediated by the ACC (Van Veen & Carter, 2002) that take place before response execution (Folstein & Van Petten, 2008; Yeung et al., 2004). We also found that the amplitude of the N2 was largest for trials following incongruent ones, which parallels the results obtained at the behavioral level and is consistent with an explanation in terms of conflict monitoring (Yeung et al., 2004). As mentioned above, a close relationship has been suggested between behavioral conflict data and amplitude modulations in the N2 potential (Folstein & Van Petten, 2008; Yeung et al., 2004). It is therefore reasonable to assume that the 'more careful' mode promoted by the resolution of previous incongruences remains in the current trial; this seems to lead to a deeper evaluation of online responses based on those already executed regardless of the current trial type. By contrast, we did not observe any N2 modulation in the affective task. Although surprising, this result agrees with those recently obtained by Ruz, Madrid and Tudela (2012). These authors employed a novel paradigm to study the effect of conflictive emotions during interpersonal interactions, in which participants had to judge the emotions displayed by their partners in an economic game. Although emotional conflict was found at the behavioral level, this effect did not modulate the amplitude of the N2 potential. On the other hand, other previous studies have argued that the emotional nature of the materials modulates the N2 (Kanske & Kotz, 2010b, 2011a). These authors used SRC tasks in which participants had to respond according to the cognitive nature of target stimuli displayed upon an affective background. Their results showed larger amplitudes for the affective than for the neutral background along with a conflict effect only for affective materials (Kanske & Kotz, 2010b, 2011a). As these authors themselves explain, their tasks evaluated the influence of affective materials over the resolution of 'cognitive' conflict and not affective conflict resolution *per se* (Kanske & Kotz, 2011a). Thus, whereas in the current experiment and also in Ruz et al. (2012) participants performed explicit affective evaluations, the results obtained by Kanske and Kotz (2010b, 2011a,b; see also Gootjes, Coppens, Zwaan, Franken & Van Strien, 2011; Ma, Wang, Wang, Wang & Wang, 2010; Van Hooff, Dietz, Sharma & Bowman, 2008) may reflect the influence of the mood induced by the emotional material over cognitive conflict resolution, rather than emotional task-conflict. Further experiments directly comparing explicit and implicit affective evaluations could help clarify this issue.

The last potential explored in our study was the P3 wave, which was divided into three sections ($P3_1$, $P3_2$ and $P3_3$) according to the maps revealed by the topographic analysis. Interestingly, the decision and response processes associated to the P3 (Luck, 2005; Polich, 2006) interacted with the cognitive or affective nature of the task in different epochs, whereas the nature of the previous trial modulated the potential for both tasks. The first part of the P3 displayed larger amplitudes for current congruent trials than for incongruent trials, but only in the cognitive task. This result has been previously obtained in similar studies on cognitive conflict (Neuhaus et al., 2010; Valle-Inclán, 1996a). It has been established that stimuli that are most difficult to resolve usually show lower amplitudes in the P300 than easier ones (Johnson, 1986; Valle-Inclán, 1996b). This evidence is consistent with the inhibition hypothesis proposed by Polich (2006), who postulated the existence of mechanisms involved in the inhibition of 'extraneous' task events underlying P3 modulations. Thus, a correct or higher inhibition would allow a greater available amount of attentional resources for target processing. Hence, incongruent trials that are more difficult to resolve also leave less cognitive resources available and, as a consequence, the incongruent condition generates smaller amplitudes in the P3.

By contrast, this potential was modulated by the current emotional conflict in the last part of the P300. Previous research suggests a special sensitivity of the P3 to the affective nature of the stimuli (Cano, Class & Polich, 2009; Gootjes et al., 2011; Olofsson et al., 2008; Polich, 2006; Wu & Zhou, 2009). This modulation also resembles the late positive complex/potential (LPC/LPP; Olofsson et al., 2008) commonly studied in emotion paradigms (see Cuthbert, Schupp, Bradley, Birbaumer & Lang, 200; Schupp, Junghöfer, Weike & Hamm, 2004). In this epoch, incongruent trials displayed a P3 of larger amplitude than congruent ones, which reverses the pattern observed for the cognitive task in the first P3 section. Chiu et al. (2008) obtained amplitude results that were similar to our findings using an emotional Go/NoGo task. These amplitude variations may reflect a special discriminability of the emotional materials due to their motivational salience (Chiu et al., 2008; Polich, 2006). In a similar way, Zhang, Kong and Jiang (2012), found a valence congruency effect reflected in larger amplitudes for the incongruent condition in the LPC/LPP. Zhang et al. (2012) explained their results as reflecting a special modulation of the affective stimuli at later processing stages due to the influence of valence over semantic categorization operations. Finally, in the first two sections of the P3, trials following a congruent trial had a larger amplitude regardless of the type of task, whereas the same modulation was restricted to the cognitive version in the third part of the P3 potential. Thus, it is reasonable to argue that when the previous trial demanded fewer cognitive resources, as was the case in the congruent situation, there were more resources available for the following trials. This could have generated an enhancement of all types of subsequent trials, reflected in larger P3 amplitudes. The absence of this effect in the third part of the wave for the emotional task may be explained by a greater engagement of resources owing to the special salience of affective stimuli, as mentioned above.

Despite all this evidence, one intriguing result is the temporal mismatch between the average speed of responses and the latency of the emotional conflict effect in the third section of the P3 potential. Note that this modulation occurs in a temporal window in which the average response has already taken place, which could be interpreted to mean that the behavioral emotional conflict has no reflection in the ERP data. An explanation for this could be that the modulation observed in the P3 was partly a cortical reflection of conflict-related computations that took place earlier in time (before response execution) in brain regions (i.e. subcortical structures) that are not well-captured by electrophysiological scalp measurements, such as the amygdala (Etkin, Egner, Peraza, Kandel & Hirsch, 2006; Sabatinelli, Keil, Franka & Lang, in press). This subcortical structure, which belongs to the limbic system, seems to have a reciprocal communication with the ACC during affective conflict processing (Etkin et al., 2006). In line with this evidence, similar subcortical areas seem to be engaged in emotion evaluation and their activity may be reflected in the P3 potential (Linden, 2005) and/or the LPC/LPP (Sabatinelli et al., in press). Therefore, there may exist earlier conflict detection and/or control adjustments for emotional conflict that we are not able to observe with our current methodology and paradigm. On the other hand, the P1 and N170 potentials already track emotional conflict in interaction with the previous context. Similarly, the first two parts of the P3 potential were modulated by the previous context regardless of the type of task. In short, these effects suggest that both types of information processing already respond to the conflictive nature of the materials at early stages and even at the beginning of the P3.

To our knowledge, the current study is the first to conduct a direct comparison of the time course of cognitive and affective conflict processing using EEG recordings in a task including an explicit evaluation of both kinds of materials. Although we found a common conflict effect on behavioral data, the ERP analysis showed both shared and dissociable effects for both tasks. This suggests that the first levels of information processing, mainly perceptual analyses, deal with conflict through the same mechanisms. By contrast, later stages involved in cognitive control and response planning diverge according to the nature of the task requirements generating the conflict. Further studies could be aimed at exploring these commonalities and divergences with different types of materials and probing whether similar effects are found in different settings, such as the social realm.

3.5. References

Aranda, C., Madrid, E., Tudela, P., & Ruz, M. (2010). Category expectations: a differential modulation of the N170 potential for faces and words. *Neurophychologia*, *48*, 4038–4045.

Baines, S., Ruz, M., Rao, A. L., Denison, R., & Nobre, A. C. (2011). Modulation of neural activity by motivational and spatial biases. *Neurophychologia*, *49*, 2489–2497.

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

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring vs. selection-for-action in anterior cingulate cortex. *Nature*, *402*, 179–181.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*, 215–222.

Cano, M. E., Class, Q. A., & Polich, J. (2009). Affective valence, stimulus attributes, and P300: color vs. black/white and normal vs. scrambled images. *International Journal of Psychophysiology*, *71*, 17–24.

Chiu, P. H., Holmes, A. J., & Pizzagalli, D. A. (2008). Dissociable recruitment of rostral anterior cingulate and inferior frontal cortex in emotional response inhibition. *NeuroImage*, *42*, 988–997.

Cohen, J. D., Botvinick, M., & Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: Who's in control? *Nature Neuroscience*, *3*, 421–423.

Cojan, Y., Archimi, A., Cheseaux, N., Waber, L., & Vuilleumier, P. (2013). Time-course of motor inhibition during hypnotic paralysis: EEG topographical and source analysis. *Cortex 49, 423–436*.

Correa, A., Rao, A., & Nobre, A. C. (2009). Anticipating conflict facilitates controlled stimulus-response selection. *Journal of Cognitive Neuroscience*, *8*, 1461–1472.

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, *52*, 95–111.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective* & *Behavioral Neuroscience*, *7*, 380–390.

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences*, *12*, 374–380.

Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2008). Dissociable neural systems resolve conflict from emotional vs. nonemotional distracters. *Cerebral Cortex*, *18*, 1475–1484.

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, *8*, 1784–1790.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a non search task. *Perception and Psychophysics*, *16*, 143–149. Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron*, *51*, 871–882.

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.

Foxe, J. J., Murray, M. M., & Javitt, D. C. (2005). Filling-in in schizophrenia: A high-density electrical mapping and source-analysis investigation of illusory contour processing. *Cerebral Cortex*, *15*, 1914–1927.

Funes, M. J., Lupiáñez, J., & Humphreys, G. J. (2010). Analyzing the generality of conflict adaptation effects. *Journal of Experimental Psychology: American Psychological Association Human Perception and Performance*, *36*, 147–161.

Gootjes, L., Coppens, L. C., Zwaan, R. A., Franken, I. H. A., & Van Strien, J. W. (2011). Effects of recent word exposure on emotion—word stroop interference: An ERP study. *International Journal of Psychophysiology*, *79*, 356–363.

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*, 480–506.

Grundy, J. G., Benarroch, M. F. F., Woodward, T. S., Metzak, P. D., Whitman, J. C., & Shedden, J. M. (2011). The bivalency effect in task switching: Event-related potentials. *Human Brain Mapping 34, 999–1012*.

Heil, M., Osman, A., Wiegelmann, J., Rolke, B., & Hennighausen, E. (2000). N200 in the Eriksen-task: Inhibitory executive processes? *Journal of Psychophysiology*, *14*, 218–225.

Hillyard, S. A. (2009). Imaging techniques: Event-related potentials (ERPs) and cognitive processing. In: L. Squire (Ed.), *New encyclopedia of neuroscience (Vol. 4, pp. 13-18)*. Oxford: Academic Press.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society: Biological Sciences*, *353*, 1257–1270.

Johnson, R. (1986). A triarchic model of P300 amplitude. *Psychophysiology*, *23*, 367–384.

Kanske, P., & Kotz, S. A. (2010a). Emotion triggers executive attention: ACC and amygdala responses to emotional words in a conflict task. *Human Brain Mapping*, *32*, 198–208.

Kanske, P., & Kotz, S. A. (2010b). Modulation of early conflict processing: N200 responses to emotional words in a flanker task. *Neuropsychologia*, *48*, 3661–3664.

Kanske, P., & Kotz, S. A. (2011a). Conflict processing is modulated by positive emotion: ERP data from a flanker task. *Behavioural Brain Research*, *219*, 382–386.

Kanske, P., & Kotz, S. A. (2011b). Positive emotion speeds up conflict processing: ERP responses in an auditory Simon task. *Biological Psychology*, *87*, 122–127.

Kopp, B., Rist, F., & Mattler, U. (1996). N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology*, *33*, 282–294.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility -A model and taxonomy. *Psychological Review*, *97*, 253–270.

Lehmann, D. (1987). Principles of spatial analyses. In: A. S. Gevins, & A. Remond (Eds.), *Handbook of electroencephalography and clinical neurophysiology: Methods of analysis of brain electrical and magnetic signals (pp. 309-354)*. Amsterdam: Elsevier.

Linden, D. E. (2005). The P300: Where in the brain is it produced and what does it tell us? *Neuroscientist*, *11*, 563–576.

Luck, S. (2005). *An introduction to the event-related potential technique (Vol. 1, pp. 1-49).* Cambridge, MA: The MIT Press.

Luck, S. J., Fan, S., & Hillyard, S. A. (1993). Attention-related modulation of sensory- evoked brain activity in a visual search task. *Journal of Cognitive Neuroscience*, *5*, 188–195.

Luck, S., Woodman, G., & Vogel, E. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, *4*, 432–440.

Ma, Q., Wang, K., Wang, X., & Wang, L. (2010). The influence of negative emotion on brand extension as reflected by the change of N2: A preliminary study. *Neuroscience Letters*, *485*, 237–240.

MacDonald, A. W., III, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*, 1835–1838.

Meeren, H. K. M., Van Heijnsbergen, C. C. R. J., & Gelder, B. L. M. F. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 16518–16523.

Michel, C. M., Henggeler, B., & Lehmann, D. (1992). 42-channel potential map series to visual contrast and stereo stimuli: Perceptual and cognitive event-related segments. *International Journal of Psychophysiology*, *12*, 133–145.

Morand, S. M., Harvey, M., & Grosbras, M. Parieto-occipital cortex shows early target selection to faces in a reflexive orienting task. Cerebral Cortex, http://dx.doi. org/10.1093/cercor/bhs368, in press.

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: A step- by-step tutorial review. *Brain Topography*, *20*, 249–264.

Neuhaus, A. H., Urbanek, C., Opgen-Rhein, C., Hahn, E., Ta, T. M. T., Koehler, S., Gross, M., & Dettling, M. (2010). Event-related potentials associated with Attention Network Test. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *76*, 72–79.

Nobre, A. C., Griffin, I. C., & Rao, A. (2008). Spatial attention can bias search in visual short-term memory. *Frontiers in Human Neuroscience*, 1–9.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In: R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation (Vol. 1, pp. 1—18).* New York, NY: Plenum Press. Ochsner, K., Hughes, B., Robertson, E., Cooper, J., & Gabrieli, J. (2008). Neural systems supporting the control of affective and cognitive conflicts. *Journal of Cognitive Neuroscience*, *21*, 1841–1854.

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology*, *77*, 247–265.

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: Model estimation and validation. *IEEE Transactions on Bio-Medical Engineering*, *42*, 658–665.

Perrin, Pernier, Bertrand, J., & Echallier, J. F, O. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*, 184–187.

Polich, J. (2006). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*, 2128–2148.

Posner, M. I., & Fan, J. (2004). Attention as an organ system. In: J. R. Pomerantz, & M. C. Crair (Eds.), *Topics in integrative neuroscience: From cells to cognition*. Cam- bridge, UK: Cambridge UniversityPress.

Redondo, J., Fraga, I., Padrón, I., & Comesaña, M. (2007). The Spanish adaptation of ANEW (Affective Norms for English Words). *Behavior research methods*, *39*, 600–605.

Ruz, M., & Nobre, A. C. (2008a). Attention modulates initial stages of visual word processing. *Journal of Cognitive Neuroscience*, *20*, 1727–1736.

Ruz, M., & Nobre, A. C. (2008b). Dissociable top-down anticipatory neural states for different linguistic dimensions. *Neuropsychologia*, *46*, 1151–1160.

Ruz, M., Madrid, E., & Tudela, P. (2013). Interactions between perceived emotions and executive attention in an interpersonal game. *Social Cognitive & Affective Neuroscience*. 8, 838-44.

Sabatinelli, D., Keil, A., Franka, D. W., & Lang, P. J. 2013 Emotional perception: Correspondence of early and late event-related potentials with cortical and subcortical functional MRI. *Biological Psychology*, 92 (3), 513–519.

Scerif, G., Worden, M. S., Davidson, M., Seiger, L., & Casey, B. J. (2006). Context modulates early stimulus processing when resolving stimulus-response conflict. *Journal of Cognitive Neuroscience*, *18*, 781–792.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. *Pittsburgh*: *Psychology Software Tools Inc*.

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2004). The selective processing of briefly presented affective pictures: An ERP analysis. *Psychophysiology*, *41*, 441–449.

Snee, R. (1977). Validation of regression models: Methods and examples. *Technometrics*, *19*, 415–428.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Tucker, D. (1994). Spatial sampling of head electrical fields: The geodesic sensor net. *Electroencephalography and Clinical Neurophysiology*, *87*, 154–163.

Valle-Inclán, F. (1996a). The locus of interference in the Simon effect: An ERP study.

Biological Psychology, *43*, 147–162.

Valle-Inclán, F. (1996b). The Simon effect and its reversal studied with event-related potentials. International Journal of Psychophysiology, 23, 41—53.

Van Hooff, J., Dietz, K., Sharma, D., & Bowman, H. (2008). Neural correlates of intrusion of emotion words in a modified Stroop task. *International Journal of Psychophysiology*, *67*, 23–34.

Van Veen, V., & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, *14*, 593–602.

Verguts, T., Notebaert, W., Kunde, W., & Wühr, P. (2011). Post-conflict slowing: Cognitive adaptation after conflict processing. *Psychonomic Bulletin* & *Review*, *18*, 76–82.

Whalen, P., Bush, G., McNally, R. J., Wilhelm, S., McInerney, S. C., Jenike, M. A., & Rauch, S. L. (1998). The emotional counting Stroop paradigm: A functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biological Psychiatry*, *44*, 1219–1228.

Woodward, T. S., Meier, B., Tipper, C., & Graf, P. (2003). Bivalency is costly: Bivalent stimuli elicit cautious responding. *Journal of Experimental Psychology*, *50*, 233–238.

Woodward, T. S., Metzak, P. D., Meier, B., & Holroyd, C. B. (2008). Anterior cingulate cortex signals the requirement to break inertia when switching tasks: A study of the bivalency effect. *NeuroImage*, *40*, 1311–1318.

Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, *1286*, 114–122.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931–959.

Zhang, Q., Kong, L., & Jiang, Y. (2012). The interaction of arousal and valence in affective priming: Behavioral and electrophysiological evidence. *Brain Research*, *1474*, 60–72.

Zhu, H. R., Zhang, H. J., Wu, T. T., Luo, W. B., & Luo, Y. J. (2010). Emotional conflict occurs at an early stage: Evidence from the emotional face—word stroop task. *Neuroscience Letters*, *478*, 1–4.

Capítulo 4: Serie Experimental II. Ignoring facial emotion expressions does not eliminate their influence on cooperation decisions

El contenido de este capítulo ha sido publicado como Alguacil, S., Tudela, P., & Ruz, M. (2015). Ignoring facial emotion expressions does not eliminate their influence on cooperation decisions. *Psicológica*, 36(2), 309-335.

Abstract

Whereas the automaticity of emotion processing has been investigated in several cognitive domains, its mandatory influence on cooperative decision-making is still unexplored. We employed an interference-task to evaluate whether explicit instructions to ignore the emotions of others during alleged interpersonal interactions override their behavioral effects. Participants played a Trust Game multiple times with eight cooperative or non-cooperative partners, who displayed facial expressions of happiness or anger. Emotions were non-predictive regarding the partners' cooperation. In Experiments 1 and 2 participants were explicitly asked to ignore the emotions, and the uncertainty about the partners' behavior varied. We found an effect of emotional interference; whereas happy partners speeded cooperative decisions, angry ones speeded non-cooperative choices. This was replicated in Experiment 3, where the request of ignoring emotions was removed. Our results show the inevitable influence of the emotional displays of others during cooperation decisions, which fits with theories that contend for a tight link between emotions and social context.

4.1. Introduction

We live in environments crowded with multiple stimuli. For a successful navigation that fulfills current goals and allows optimal decision-making, humans rely on selection mechanisms that enhance the processing of relevant information whilst keeping distracting events out of the attentional focus (Egner, Etkin, Gale & Hirsch, 2008; Posner, 2011; Squire, Noudoost, Schafer & Moore, 2013). However, certain types of salient stimuli escape control mechanisms and reach decision and response stages even when they are explicitly ignored (e.g. Pourtois, Schettino & Vuilleumier, 2013). This is taken as proof that such information is processed in an automatic or mandatory, non-volitional fashion. Affective events are among the most salient types of stimuli, and they have the potential to modulate decision-making. The studies reported in this paper evaluate whether the influence of emotional facial expressions on alleged social interactions can be controlled by explicit instructions, and the temporal distribution of these effects.

Some theories contend that the constant interplay between internal and external affective environments impacts how emotional states, beliefs, behaviors and decisions arise (Augoustinos, Walker & Donaghue, 2006; Fischer & Manstead, 2008). In line with this, social context seems essential for emotions to have meaning and function (Barrett, 2012; Keltner & Gross, 1999; Parkinson, Fischer & Manstead, 2004; Parrott & Schulkin, 1993). According to Fischer and Manstead (2008), emotions are decisive elements for social survival and help us deal with the challenges that interpersonal environments pose. They facilitate adaptation by readying, automatically and efficiently, context-appropriate behavioral responses, and making the individual more likely to engage in efficient action (Frijda, 1988; Frijda & Mesquita, 1994; Tooby & Cosmides, 1990). At the *intrapersonal* level emotions provide relevant information about social events, and prepare us to act as needed (Keltner & Haidt, 1999). During *interpersonal* interactions, the emotional displays of others yield clues about their internal states and their most likely behavior (see for example, Ekman & Friesen, 2003). Thus, emotional displays help the agent gather information about the emotions, beliefs and future intentions of others, as well as their appraisal of the current situation, and serve as tools for coordinating interactions among people (Keltner & Haidt, 1999). Such interpersonal interactions depend on the affiliation and distancing functions of emotions. Whereas the former would help to form and maintain relations with others, the latter would lead the agent to differentiate from others and compete with them. In this way, emotions help to balance the social goals of cooperation and competition to an optimum compromise (Fischer & Manstead, 2008). Action tendencies of cooperation and competition are mainly built on social judgments (Forgas, 1991). These appraisals are affected by quick, automatic evaluations, which are strongly grounded on emotional information (Haidt, 2001). The accuracy and efficiency of judgments are of critical relevance for our private life and especially for an effective interpersonal behavior (Forgas, 1991).

Economic games are useful paradigms to study how emotions, in addition to other factors, impact our behavior when cooperation and competition tendencies emerge during interpersonal interactions (Camerer, 2003). Emotions have an impact on decisions during negotiation scenarios (e.g. Kopelman, Gewurz & Sacharin, 2007). For example, the experience of positive affect is associated with a higher number of cooperative behaviors and joint gain seeking, whereas negative feelings have the opposite effect (Forgas, 1998; Kopelman et al., 2007). The negative reactions generated by unfair offers in the Ultimatum Game (UG) provoke punishment behaviors, even when people have to sacrifice their future benefits (Sanfey, 2009; Sanfey, Rilling, Aronson, Nystrom & Cohen, 2003).

The modulation of cooperation and competition tendencies as a consequence of emotional information is also crucial at the interpersonal level. Communicated anger can promote different patterns of behavior. For example, in the UG, expressions of anger can increase the amount of rejections of offers (Kopelman, Rosette & Thompson, 2006). But they also foment generous behavior in participants if they are encouraged to consider the opponent's emotions (see for example Van Kleef, De Dreu & Manstead, 2004) or such emotional states are directly linked to the offer of the proposer (Lelieveld, Van Dijk, Van Beest, Steinel & Van Kleef, 2011). In the Trust Game (TG), on the other hand, happy facial expressions generate higher levels of initial and sustained trust even in contexts in which they lack predictive value (Eckel & Wilson, 2003; Scharlemann, Eckel, Kacelnik, & Wilson, 2001; Tortosa, Strizhko, Capizzi & Ruz, 2013; see also Averbeck & Duchaine, 2009; Campellone & Kring, 2013).

Although the field of economic bargaining has mostly stressed the role of strategic factors on decisions (Camerer, 2003; Lee & Harris, 2013), judgments of trust also seem to be guided by the *non-volitional* processing of facial features (Oosterhof & Todorov, 2008). Several studies have shown that facial dimensions guide judgments of trustworthiness of unfamiliar people (e.g. Todorov, Baron & Oosterhof, 2008; Todorov, Mandisodza, Goren & Hall, 2005). The evaluation of emotionally neutral faces in terms of trustworthiness seems to be an over-generalization of functionally adaptive systems for detecting emotions in others, which appear to be based on basic facial features that resemble emotional expressions signaling approach/avoidance tendencies (Todorov et al., 2008). In this over-generalization model, faces at the negative extreme of the trustworthiness dimension are related to the expression of anger, whilst those at the positive extreme are linked to happiness features (Todorov, Said, Engell & Oosterhof, 2008). Angry faces arise avoidance tendencies whereas happy faces signal the opposite (Todorov et al., 2008). In this line, structural features in faces resembling happy expressions are associated with judgments of trustworthiness whereas those characteristics similar to anger are linked to judgments of untrustworthiness (Todorov et al., 2008).

The effect of emotions on our social judgments can be controlled (e.g. Ruz & Tudela, 2011; Satpute, Badre & Ochsner, 2013), although the extent of this control has not yet been established. We are endowed with the ability to flexibly regulate our emotions, and on many occasions we are able to effectively induce or suppress emotional reactions through a variety of tactics (Gross, 1999). We can thus modulate the impact of emotions, felt and perceived, when interacting with other people in bargaining scenarios (Kopelman et al., 2006). This ability can be particularly useful during competition, where other people may conceal or fake their true dispositions (Dawkins & Krebs, 1978; Fridlund, 1994; Krebs & Davies, 2009). However, emotion regulation demands cognitive control to manage the behavioral reactions that emotions from others generate in us, and to avoid being misled. An example of the deployment of cognitive control in interpersonal scenarios comes from the use of situations in which emotions conflict with the expectations they generate, which leads to a marked increase in decision times (Ruz, Madrid & Tudela, 2013; Ruz & Tudela, 2011). Previous studies on this respect, however, instructed participants to use the emotional expression of the partner as a cue to predict their future behavior, and thus our capacity to actively ignore emotions and their ability to intrude our decisions has not yet been tested.

Hence, the extent of top-down control guided by explicit instructions over the approach-avoidance reactions that emotions generate during interpersonal interactions is an open question. In our series of experiments, we investigated whether the influence of fully non-informative emotions on interpersonal decisions could be completely eliminated by explicitly instructing participants to ignore them. Along the series, we employed an adaptation of the Trust Game (Berg, Dickhaut & McCabe, 1995) where participants decided if sharing or not their money with trustworthy and untrustworthy partners in a multi-round setting. Game partners displayed emotional expressions that participants were asked to ignore. We manipulated the type of partner, their emotional display, the validity of the personal information, and the explicit instruction of ignoring the affective information.

In addition, we explored how the putative effect of emotions on decisions changed depending on the speed of responses, following the model proposed by Ridderinkhof (2002). This model poses that the amount of mental resources and the time taken to respond influence the mental strategies in operation during conflict resolution. According to the activation–suppression model, stimuli directly linked to responses (S-R) do not require a large amount of mental resources, as they take place in a mostly ballistic manner. However, when such S-R connections lead to incorrect responses, they are followed by the online engagement and implementation of suppression mechanisms of cognitive control (Ridderinkhof, 2002). By this logic, more automatic reactions would occur at the shortest response times whereas controlled responses would take place at the slowest RTs. Thus, automatic and controlled strategies would influence the distribution of behavioral performance (i.e. accuracy and reaction times, RTs). The strategy to implement a distributional analysis is based on the division of the full range of data into several temporal intervals, which are named 'bins'. These 'bins' are included in the general statistical analysis with the aim of obtaining a detailed description of the influence of selective suppression mechanisms on choices made at different speeds (see Ridderinkhof, 2002; Ridderinkhof, van den Wildenberg, Wijnen & Burle, 2004, for a similar strategy).

Given the biological relevance of affective stimuli (LeDoux, 2001), their rapid evaluation and privileged processing (Barrett & Bliss-Moreau, 2009; Reeck & Egner, 2011), we expected to find an effect of conflict on decision times (Ruz & Tudela, 2011), driven by the action tendencies generated by the ignored emotions, in the social setting. Based on the over-generalization of trustworthiness features theory by Todorov (Todorov et al., 2008), and in line with previous studies in the same field (see for example Ruz & Tudela, 2011), we employed expressions of happiness and anger in order to maximize the emotions' avoidance-approach tendencies aroused by the partners' facial expressions. These two emotions also share characteristics, such as transmitting an elevated sense of certainty and personal control while keeping opposite valences (see for example Han, Lerner & Keltner, 2007; Lerner & Keltner, 2000). We hypothesized that happy expressions of partners would generate trustworthy evaluations in participants and thus expectations of cooperation, whereas the opposite would occur for angry facial expressions. We expected that these expectations would be reflected on the speed of choices made in relation to cooperative vs. non-cooperative partners in an interactive fashion. Happy cooperative and angry non-cooperative partners would generate faster choices than angry cooperative and happy non-cooperative ones. Also, given the hypothesized automatic nature of this effect, we predicted that its size would be larger in the fast decisions (bins), compared to the slow decisions (bins), as the cognitive control required for its suppression would demand additional resources and time to operate (Ridderinkhof, 2002; Ridderinkhof et al., 2004).

4.2. Experiment 1

4.2.1. Participants

Twenty-six right-handed healthy volunteers (13 men, mean age: 23), with normal or corrected-to-normal vision, were recruited from the University of Granada and received course credits in exchange for their participation. They all signed a consent form approved by the local Ethics Committee.

4.2.2. Stimuli, Apparatus and Procedure

Participants played the Trust Game multiple times with 8 different partners. They were instructed that in every trial they had to decide whether to share or not a sum of fictional money (€1; received at the beginning of each game round) with a partner. If they chose to share, the amount was multiplied by five and passed to the partner, who decided next whether to cooperate or not with the participant. If s/he reciprocated the participant's trust, each would earn €2.5. However, if the partner did not cooperate, s/he would get €5 and the participant €0. The participant's goal was to achieve as much money as possible with each partner. In addition, they were instructed about the cooperative or non-cooperative patterns of the partners and about their identity. Participants were also told to retain and learn the behavioral pattern of each of their partners and use this information during the experimental task. Half of them most likely cooperated with them and the other half most likely did not. This information was valid in 70% of the trials for each partner, and invalid in the remaining 30%. That is, trustworthy partners cooperated on 70% of the occasions, whereas untrustworthy ones only did so on 30% of the exchanges. Hence, the identity of the partner operated as a trustworthiness cue that participants could use to guide their decisions. In addition, participants were told that partners would display happy and angry facial emotional expressions, but that they were not relevant to the partner's choices or any other aspect of the game, and thus they had to be ignored. Participants were instructed to respond as quickly as possible, to enhance the sensitivity of reaction times to the manipulations. If their response was too slow (> 1500 ms) a text reminded them to respond faster. This reminder did not appear for responses shorter than 1500 ms. Afterwards, an asterisk or a hash symbol (either purple or blue) provided feedback. Participants were instructed about the meaning of these symbols, which informed about the decision that the partner had made (in trials in which the participant cooperated) or would have made (in trials in which the participant decided not to cooperate).

Sixteen facial stimuli (8 identities, 4 females and 4 males) displaying happy or angry (50%) emotional expressions were taken from the NimStim set (Tottenham, Tanaka, Leon, McCarry, Nurse, Hare et al., 2009). These faces were used as partners in the game. Emotional expressions were orthogonal to the partner types and their cooperation rates; hence, emotions were not predictive of the partners' cooperative behavior.

The trustworthiness of the identities (cooperative or non-cooperative), the association between hand and response key and the feedback symbols (and their ink color) were fully counterbalanced across participants. The task was created and displayed using E-Prime 2.0 Professional software (Schneider, Eschman and Zuccolotto, 2002). Trials were presented in a random order, and the stimuli were displayed on a silver background. Each trial comprised the following sequence (see Figure 1): A symbol of 1 Euro (**€;** 1.9° x 2.39°) displayed for 200 ms at the center of the screen, followed by a fixation point (+; 0.57°) of 500 ms on average (random 250-750 ms). Next, the face of the partner for the trial appeared during 1500 ms (6.20° x 7.15° on average), followed by another fixation point identical to the first one. Afterwards a feedback symbol ('*', 0.67° x 0.7°, or '#', 0.57° x 0.95°;

displayed in blue or purple ink) was displayed for1000 ms, and finally a third fixation point with the same characteristics as the previous ones ended the trial. Participants were instructed to respond during the time the partner's face was present on the screen by pressing one of two buttons with their left or right index fingers. On average, a trial lasted 4200 ms.

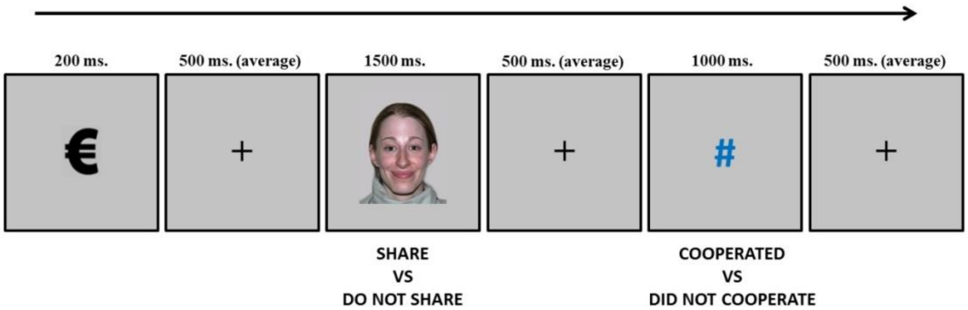

Figure 1: Sequence of trial events in all experiments.

Before the main task, participants performed a training block of 40 trials to familiarize themselves with the procedure. The stimulus set and instructions about face and feedback were the same in practice and experimental phases. The only difference was that, during practice trials, the cooperation of the partners was 100% consistent, to increase learning about their cooperation tendencies. The experiment was composed of 4 blocks of 80 trials each (320 in total), separated by 3 brief breaks. The approximate duration of the task was 22 minutes. When the task ended, participants completed a short questionnaire (10 point-Likert scale ranging from 0 to 10) in which they indicated the approximate percentage of cooperation of each of their game partners. All participants included were able to indicate the cooperation patterns of all their partners.

4.2.3. Data analysis

Following procedures already published (e.g. Ridderinkhof et al., 2004), the percentage of cooperation rates and the mean RTs were rank-ordered per condition and split up in three equivalent RT speed tercils (bins), and the average in each condition submitted to an ANOVA with the factors displayed Emotion (happy, angry), Partner Trustworthiness (cooperative, non-cooperative) and Bin (tercils 1, 2, 3). Trials without a response (i.e. sharing or not the money) were eliminated from the bin selection and subsequent analyses. All of the remaining trials, regardless of whether the choice (sharing or not) corresponded to the type of partner (cooperative, non-cooperative), were included in the analyses[3]. The inclusion of the bin factor afforded the consideration of additional information in our design, which allowed us to explore the automatic vs. more controlled effect of emotions on social decisions. This is also of interest to compare the current results with similar results from non-social interference paradigms (Ridderinkhof, 2002; Ridderinkhof et al., 2004). The removal of the bin factor did not change any

[3] Including only choices congruent with the type of partner (cooperating only with cooperative partners and not doing so with non-cooperative ones) did not change the general pattern of results.

of the other findings in the studies reported below. In any case, we kept the bin factor in the design given the additional information this variable provides. In addition, in all Experiments we performed an additional ANOVA including the factor of Block, to account for potential practice or learning on bin effects (given that slower responses may take place in earlier initial trials compared to late trials in the experiments). We did not observe any main effect of block or interaction with the Bin factor, which rules out interpretations of the Bin in terms of practice with the task.

4.2.4. Results

Only trials in which participants did not respond were discarded from the analyses (1.14 %). On average, participants cooperated on 52 % of the trials. The ANOVA of the mean cooperation rates showed a main effect of Emotion, F (1, 25) = 5.99, $p < .05$, $\eta p2 = .19$, as cooperation rates were higher for partners displaying happy, M = 55 %, SD = 6 %, than angry, M = 49 %, SD = 7 %, emotional expressions. Partner Trustworthiness was also significant, F (1, 25) = 258.70, $p < .001$, $\eta p2 = .91$. Participants shared their money more frequently with cooperative, M = 89 %, SD = 12 %, than with non-cooperative, M = 15 %, SD = 11 %, partners. In addition, there was an interaction between the three factors in the design, F (1, 25) = 4.74, $p < .05$, $\eta p2 = .28$. This was because whereas the interaction between Emotion and

Partner Trustworthiness was not significant in the first two bins, all *ps* > .118, it reached significance in the third one, F (1, 25) = 9.96, $p < .01$, $\eta p2 = .28$. Subsequent analyses in this third bin showed an effect of the emotional display close to significance levels for cooperative partners, F (1, 25) = 3.90, $p = .059$, $\eta p2 = .13$, and significant for non-cooperative partners, F (1, 25) = 17.07, $p < .001$, $\eta p2 = .41$. Participants shared more money with happy than with angry non-cooperative partners, *M* = 18 %, *SD* = 14 % vs. *M* = 9 %, *SD* = 8 %; the tendency for cooperative partners was in the same direction.

The ANOVA of the mean RTs revealed a main effect of Emotion, F (1, 25) = 6.68, $p < .05$, $\eta p2 = .21$, as participants' responses were faster for happy, *M* = 596.82, *SD* = 118.61, than for angry partners, *M* = 599.49, *SD* = 117.06. Partners' Trustworthiness was also significant, F (1, 25) = 9.83, $p < .01$, $\eta p2 = .28$, as decisions were slower for non-cooperative, *M* = 613.36, *SD* = 124.47, than for cooperative partners, *M* = 582.95, *SD* = 113.37. Crucially, these two factors interacted, F (1, 25) = 4.52, $p <.05$, $\eta p2 = .15$. Decisions were slower for angry, *M* = 595.71, *SD* = 119.03, than for happy, *M* = 570.26, *SD* = 109.05, cooperative partners, F (1, 25) = 25.46, $p < .001$, $\eta p2 = .50$, whereas this pattern was reversed for non-cooperative ones, F (1, 25) = 5.03, $p < .05$, $\eta p2 = .17$, as responses were faster for angry, *M* = 603.28, *SD* = 120.61, than for happy, *M* = 623.43; *SD* = 132.11, partners (see Figure 2). The bin factor did not interact with any other variable, all *ps* > .159.

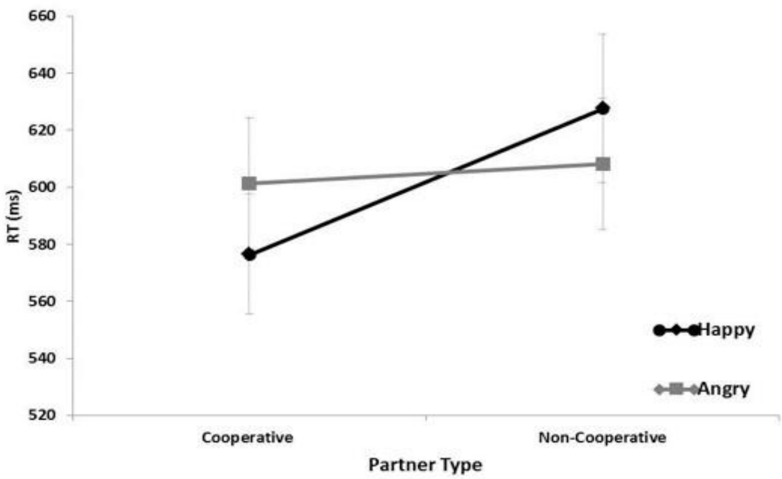

Figure 2: Interaction of Emotion and Partner Trustworthiness on the speed of choices in Experiment 1.

4.2.5. Discussion

Data from Experiment 1 support the existence of an affective bias driven by the emotions that participants were explicitly asked to ignore. In the first place, the percentage of trials in which they shared money was modulated by emotional displays, as reflected in larger concessions for happy than for angry partners. In the second place, results from this first experiment showed a clear delay in decisions due to the influence of emotional displays. On this respect, regardless of the speed of responses (i.e. the temporal moment, or bin, in which they occurred), participants took longer to make a choice when they were playing with a partner with a facial display opposite to their cooperative tendencies. In situations where participants chose whether sharing or not their money with a partner described as mostly cooperative but displaying a negative emotion, or with a happy non-cooperative partner, the decision-making process slowed down. As the emotional facial information was not predictive of the partners' cooperation rates and participants were told explicitly that it was irrelevant for the task and that they should ignore it, its influence on participants' decisions points to a failure in the full implementation of control mechanisms. In addition, the interference was equal across the whole distribution of response speeds (or bins). Thus, results suggest that affective stimuli are processed in a mandatory or non-volitional manner during the current interpersonal setting.

However, it could be the case that the uncertainty associated to the inconsistent behavior of the partners of the game (who only followed their trustworthiness pattern on 70% of the trials) led participants to pay attention to the emotions as an additional source of information, which could have generated the conflict effect observed. Contextual uncertainty entails risk and can modulate choices (see for example Gaertig, Moser, Alguacil & Ruz, 2012; Ruz, Moser & Webster, 2011). The greater the amount of perceived risk and uncertainty, the larger the influence of other contextual factors on decisions (Kopelman et al., 2007). Therefore, in Experiment 2 we increased the validity of the personal information by having partners behaving in a cooperative or non-cooperative fashion in all trials (that is, personal information was 100% valid). Thus, participants had all the information relevant for their trust choices in the identity of their partners. Obtaining an interference effect due to the emotional displays would be stronger evidence for the mandatory processing of facial emotional expressions during interpersonal choices.

4.3. Experiment 2

4.3.1. Participants

Twenty-six healthy volunteers (5 left-handed, 13 men, mean age: 23.4), with normal or corrected-to-normal vision, were recruited from the University of Granada and participated in exchange for course credits. They all signed a consent form approved by the local Ethics Committee.

4.3.2. Stimuli, Apparatus and Procedure

All details were the same as in Experiment 1 with the exception of the validity of the partner’s trustworthiness in the experimental task. In the current experiment, each partner behaved according to his/her cooperation type in all trials. That is, cooperative partners reciprocated on 100% of the game rounds and non-cooperative partners never cooperated. The ANOVAs for the cooperation rates and mean RTs were carried out including the same factors as in the Experiment 1 (Emotion, Partner Trustworthiness and Bin).

4.3.3. Results

Trials without response were eliminated from the analyses (1.02 %). Participants shared their money on 50 % of the trials. The ANOVA of the mean cooperation rates showed a main effect of Partner Trustworthiness, F (1, 25) = 2956.97, $p < .001$, $\eta p2 = .99$. Participants shared more with cooperative, M = 96 %, SD = 4 %, than with non-cooperative, M = 3 %, SD = 3%, partners. This analysis also revealed a significant interaction between Partner Trustworthiness and Bin, F (1, 25) = 16.39, $p < .001$, $\eta p2 = .58$. Cooperation rates increased along with bin for cooperative partners, F (1, 25) = 11.50, $p < .001$, $\eta p2 = .49$; participants shared more money with them during the second, M = 98 %, SD = 5 %, and third bins, M = 98 %, SD = 2 %, than during the first one, M = 92 %, SD = 9 %, $p <.001$ and $p =.001$ respectively. There were no significant differences between the second and third bins, $F < 1$. In contrast, the main effect of Bin for non-cooperative partners, F (1, 25) = 8.76, $p = .001$, $\eta p2 = .42$, was explained by a higher cooperation rate during the first bin, M = 6 %, SD = 8 %, which was larger than in the second, M = 2 %, SD = 4 %, $p <.001$, and third bins, M = 2 %, SD = 3 %, $p <.01$.

The analysis of the RTs revealed a significant interaction between Emotion and Partner Trustworthiness, F (1, 25) = 5.05, $p < .05$, $\eta p2 = .17$.

The effect of Emotion only reached significance for cooperative partners, *F* (1, 25) = 15.03, *p* = .001, *ηp2* = .38, non-cooperative, *F* < 1, as participants were faster responding to happy, *M* = 566.98, *SD* = 84.00, than to angry, *M* = 580.60, *SD* = 90.82, partners (see Figure 3). No other effect or interactions, including the bin factor, were significant, all *ps* > .102.

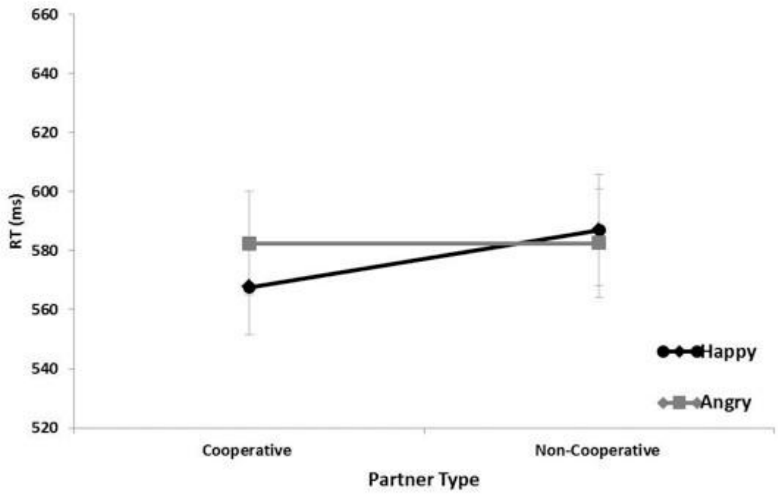

Figure 3: Interaction between Emotion and Partner Trustworthiness on the speed of choices in Experiment 2.

4.3.4. Discussion

As in Experiment 1, participants were unable to avoid processing the emotional facial expressions of their partners despite their irrelevance and the full predictability of the cooperation behavior. The increase of contextual certainty was not enough to fully remove the bias that affective information exerted on the time needed to make cooperative decisions. Unlike

Experiment 1, there were no effects of emotional information on sharing rates. Choices now were fully guided by the personal information, as reflected by larger sharing rates with cooperative than with non-cooperative partners, and this tendency was most extreme at the slowest decisions (third bin). Thus, longer time windows reflected the behavioral tendencies associated with each partner type better than the shortest ones. This suggests that the implementation of personal information about trustworthiness takes time, and is mostly effective at the longest choices.

In this experiment, the identity of the partners was fully predictive of their cooperative tendencies, and thus participants had no motive to look for additional information to guide their choices. Even though, making rational cooperative decisions in accordance with the trustworthiness of the partner took them longer when this conflicted with their facial displays, which they were told to ignore. In agreement with Experiment 1, the distributional analyses showed that the interference took place both in fast and slow responses, which suggests that suppression of affective information was equally effective in all bins. On the other hand, the reason why the interference effect was now restricted to cooperative partners is not fully clear, and we will return to it in the General Discussion section. In any case, results reinforce the argument that facial emotional expressions are processed in a mandatory way during interpersonal choices.

It could also be the case, though, that the explicit instructions to ignore the emotional expressions of the partners had the opposite effect and led participants to focus on them (e.g. Wegner, Schneider, Carter & White, 1987). Results in the field of paradoxical thought suppression suggest that trying to suppress a thought sometimes leads to a rebound, opposite effect, in which the thought receives increased processing. In the current scenario, trying to avoid paying attention to the emotional expression of the partners could have funneled their bias on behavioral responses. To test this alternative hypothesis we performed Experiment 3, where we removed the explicit instructions regarding the facial affective information of partners.

4.4. Experiment 3

4.4.1. Participants

Twenty-six right-handed healthy volunteers (13 men, mean age: 21.4), with normal or corrected-to-normal vision, were recruited from the University of Granada and participated in exchange for course credits. They all signed a consent form approved by the local Ethics Committee.

4.4.2. Stimuli, Apparatus and Procedure

All details were the same as in Experiment 2, except that participants received no instructions regarding the emotional displays that the game partners would display. Cooperation rates and mean RTs were submitted to an ANOVA including the same factors as in Experiments 1 and 2 (Emotion, Partner Trustworthiness and Bin).

4.4.3. Results

Trials without a response (i.e. sharing or not the money) were eliminated from the analyses (0.75 %). The average sharing rate was 50 %. The analysis of mean cooperation rates yielded a main effect of Partner Trustworthiness, F (1, 25) = 124.04, $p < .001$, $\eta p2$ = .83. Participants shared more with cooperative, M = 97 %, SD = 4 %, than with non-cooperative, M = 3 %, SD = 5 %, partners. This effect was qualified by its interaction with the Bin factor, F (1, 25) = 6.23, $p < .01$, $\eta p2$ = .34. Follow-up analyses showed that cooperation rates increased along with bin for cooperative partners, F (1, 25) = 3.42, $p < .05$, $\eta p2$ = .22, and decreased for non-cooperative ones, F (1, 25) = 4.56, $p < .05$, $\eta p2$ = .27. Participants shared more money with cooperative partners in the second, M = 94 %, SD = 20 %, and third, M = 94 %, SD = 19 %, p = .645, bins than in the first one, M = 90 %, SD = 20 %, both $ps < .05$. This pattern was opposite for the non-cooperative partners, as cooperation rates were higher during the first bin, M = 12 %, SD = 22 %, than during the second, M = 6 %, SD = 19 %, $p < .01$, and third ones, M = 5 %, SD = 19 %, $p < .01$.

The ANOVA of the mean RTs showed a main effect of Partner Trustworthiness, F (1, 25) = 8.32, $p < .01$, $\eta p2 = .25$; participants took longer to respond to non-cooperative, M = 567.88, SD = 66.03, than to cooperative partners, M = 552.62, SD = 66.81. This factor interacted with Emotion and Bin, F (1, 25) = 4.42, $p < .05$, $\eta p2 = .269$, and thus each bin was analyzed separately. Whereas in the first bin the only significant effect was the type of partner also mentioned above, in the second and third bins this effect was qualified by its interaction with the emotional display, F (1, 25) = 9.19, $p < .01$, $\eta p2 = .27$, and F (1, 25) = 11.45, $p < .01$, $\eta p2 = .314$, respectively. Follow-up analyses showed that in the second bin only the emotion of non-cooperative partners affected speed of decisions, Happy M = 529.51, SD = 52.33 vs. Angry, M = 518.90, SD = 52.33, F (1, 25) = 13.23, $p = .001$, $\eta p2 = .346$; cooperative, $F < 1$. However, this factor affected responses to both cooperative and non-cooperative partners in the third bin, F (1, 25) = 4.28, $p < .05$, $\eta p2 = .15$, and F (1, 25) = 12.42, $p < .01$, $\eta p2 = .33$, respectively, in opposite directions depending on whether Partner Trustworthiness was cooperative, Happy M = 619.20, SD = 79.75 vs. Angry M = 631.69, SD = 96.12, or not, Happy M = 656.01, SD = 89.26 vs. Angry M = 629.75, SD = 88.54; see Figure 4.

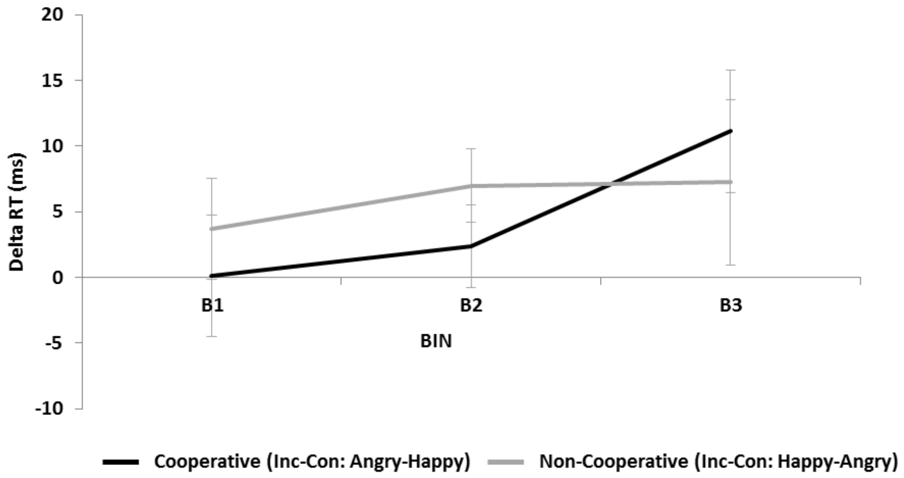

Figure 4: Delta plots for the conflict effect on RTs in Experiment 3. Display of effect size (Delta RT; incongruent-congruent conditions) as a function of response speed (RT tercil scores).

4.4.4. Discussion

In the current experiment, in which we eliminated the explicit instructions about the emotional displays of the partners, we still observed an emotional interference effect on decision times. This suggests that results in the previous experiments were not due to the unintended attention to the emotions driven by the instructions to ignore them, but to their mandatory processing during interpersonal interactions.

In contrast to Experiment 1 and 2, however, in Experiment 3 the interaction between emotion and personal information was mediated by the bin factor. At the fastest choices (first bin) only Partner Trustworthiness modulated the speed of responses, whereas this factor interacted with emotion in the other two bins. Whereas in previous studies the interaction was present in all temporal bins, results from the present experiment suggest that now the emotional conflict moved to the longer responses (second and third bins). Emotions began to affect responses to non-cooperative partners in the second bin, and they did so for both types of partners in the third one. This progressive behavioral adjustment along the three bins could reflect a natural strategy to integrate emotional information in a context in which participants did not know whether the facial expressions were relevant or not. From this perspective, longer responses would allow more time to ponder emotions as potential relevant factors, which would be reflected on their heightened effect on slower choices. We will return to this point in the following and last section.

4.5. General discussion

With the present series of experiments we show that emotions that should be ignored are still processed and influence cooperative responses in alleged social scenarios, and that this takes place in a fairly constant manner across the whole distribution of response speeds (i.e. bins). Behavior during social decision-making, mostly guided by personal identity information, was impacted by emotional irrelevant information from faces. Such effect appeared in a consistent manner across the experiments, as reflected by the interaction between the cooperative tendencies of the partners and their emotional expression, observed on the speed of choices across the 3 studies. This was also true even in a context devoid of risk where the personal information was fully predictive of the partners' behavior. In addition, the instruction of ignoring irrelevant emotional information was not the cause of the results obtained in Experiment 1 and 2. Thus, emotional contradictory information during social decision-making generates conflict that is unavoidable by cognitive control mechanisms. Our results stress the pivotal importance of emotions in interpersonal interactions and provide further support for theories that contend for a tight link between emotions and social environment (Barrett, 2012; Keltner & Gross, 1999; Parkinson et al., 2004; Parrott & Schulkin, 1993).

In our studies we measured behavioral indexes of cooperation choices and their speed, and obtained evidence of their permeability by irrelevant emotional expressions. Across the three experiments, these indices were affected by the emotional information to a different extent. The temporal profile of the effect also varied depending on the trustworthiness of the partners. Participants' cooperation rates were adjusted to the identity and the expected behavior of the partners across experiments. Crucially, cooperative choices were biased by ignored emotions when the context was uncertain (Experiment 1). This result resonates with previous evidence showing that cooperation decisions are biased by information not predictive in the task setting but that has previous associations with cooperative or non-cooperative behaviors, both in terms of social and moral information (Delgado, Frank & Phelps, 2005; Gaertig et al., 2012; Moser, Gaertig & Ruz, 2014; Ruz et al., 2011) and also emotional expressions (Campellone & Kring, 2013; Tortosa et al., 2013; Tortosa, Lupiáñez & Ruz, 2013). However, in previous studies participants were not explicitly asked to ignore the information provided, and thus the effects observed could be at least partially explained by experimenter bias. In the current study, however, participants received such instructions, and still the uncertainty of the task setting made cooperation choices vulnerable to preexisting associations between irrelevant emotional expressions and behavioral tendencies. Our results thus provide novel evidence that in ambiguous social settings, in the low certainty conditions of Experiment 1, emotions that we actively try to ignore bias our cooperation choices in an expression-congruent manner. This bias on choices disappeared on Experiments 2 and 3. One compelling approach that could help to explain the presence of emotional bias on uncertain choices is the Affect Infusion Model by Forgas (AIM; 1994; 1995). Although this theory is mainly oriented to intrapersonal affect and its consequences during social judgments, it may be relevant to the current findings. The AIM states that the more elaborate the construction of a social judgment, the greater the influence of affect. From this logic, it is reasonable to assume that in the first experiment, where the uncertainty of the context was high, the risk in the situation promoted seeking for other factors to construct a judgment and make a decision. Participants would go beyond personal information to build inferences about the partners' affective state based on prior knowledge (Forgas, 1995). Under the same rationale, the reduction of contextual uncertainty in the other two experiments could have helped participants to simplify the process and, as consequence, their choices would have been be guided mostly by a preexisting motivational goal (i.e. maximize their benefit by employing personal information) without the inclusion of other irrelevant elements in the actual choices (i.e. facial emotional expressions).

In contrast to choices, the speed of responses was affected by emotions in the three experiments, regardless of the game uncertainty and also of the instructions. However whereas in the first experiment this conflict appeared for both types of partners, the reduction of the uncertainty in Experiment 2 limited the scope of the effect to cooperative partners. The reasons for this are unclear, although previous studies suggest that people rely on theory of mind processes with others who engage in cooperative behavior (Fehr & Camerer, 2007; Frith & Frith, 2007; McCabe, Houser, Ryan, Smith & Trouard, 2001), and that these mechanisms may be partially different for cooperative and non-cooperative partners (Lissek et al., 2008;

Ruz & Tudela, 2011). It could also be the case that participants were much more focused in their strategy with cooperative partners as they were expecting larger gains from their interaction than with non-cooperative ones, and that this strengthened the effect of their emotions. On the other hand, supposing that the interference effect is driven by the incongruent happy expressions for cooperative partners and for the angry expression in the case of the non-cooperative ones, a complementary explanation for this asymmetry in the interference effect could be derived from the 'broaden theory' (Fredrickson & Branigan, 2005). This theory posits that positive and negative emotional events engage available attentional resources to a different extent. As a consequence, emotions would either narrow or broaden the attentional focus deployed to the task. Whereas negative emotions tend to narrow action tendencies, positive ones would lead to wider variability in behavior. As such, negative emotions in the case of cooperative partners would lead to a more effective interference than positive emotions for non-cooperative ones, which would lend cooperative partners more prone to interference. In any case, and regardless of the explanation, our results suggest that responses to cooperative partners are more prone to emotional interference that responses to non-cooperative ones, although further research is needed to replicate this effect and elucidate its potential causes. In any case, the removal of the request of ignoring emotions in Experiment 3 did not eliminate the conflict, which was again apparent for both types of partners at the slowest temporal window.

The analysis of the temporal profile of the decisions offered additional relevant information. Choices in situations of interpersonal uncertainty (Experiment 1) were affected by irrelevant emotional displays only in the slowest responses. In these, participants tended to cooperate more with happy than with angry partners. Removing the uncertainty of the partners' behavior in Experiments 2 and 3 eliminated the effect of emotion on actual choices. However, their cooperation tendencies interacted with bins showing that slower decisions were more adjusted to the cooperative or non-cooperative tendencies of the partners. Overall this suggests that the longer participants took to decide, the better they responded according to the type of partner. This is in line with a higher degree of control in slow response-time windows (Ridderinkhof, 2002), and provides support for the usefulness of exploring the profile of choices with different temporal distributions.

Results regarding the distribution of the emotional conflict on RT indices, on the other hand, contrast with our initial predictions following the model proposed by Ridderinkhof (2002). Whereas previous results in this line suggest that automatic interference effects are stronger when responses are fast and that control mechanisms need more time to develop and thus are mostly reflected in the slowest responses, we observed emotional conflict in both fast and slow bins (Experiments 1 and 2) or only in full in the slowest bin (Experiment 3). It is well-known that emotions carry important information for survival and are processed in a fast and efficient manner (see for example Fischer & Manstead, 2008), which leads to their prioritized processing (Vuilleumier, 2005). The finding that emotional expressions interfere with responses even in the slowest bins suggests that control mechanisms are not fully operative even in this time window. This pattern persists in Experiment 3, although the lack of explicit instructions regarding emotions in this case leads to a disappearance of the effect in the fastest responses. This temporal pattern is puzzling, and it suggests that instructions may change the speed with which affective information captures resources. Whereas most of the studies applying the Activation-Suppression model (Ridderinkhof, 2002) employ non-emotional materials in interference tasks devoid of social content (e.g. Ridderinkhof, 2002; Ridderinkhof et al., 2004), our experiments measure the extent of emotional conflict in interpersonal social contexts, which hinders comparisons. In the Stroop paradigm the nature of the specific conflict seems to be a critical element as it seems to determine which control mechanisms are implemented and how conflict is dealt with (see for example, Egner, 2008; Funes, Lupiáñez & Humphreys, 2010; Loose, Kaufmann, Tucha, Auer & Lange, 2006; Van Veen & Carter, 2005). In a complementary way, another factor that may affect the pattern of results is the rather stringent limit on response speed that our task imposed, as participants were not allowed to take longer than 1500 ms to decide. In any case, further studies should compare the temporal course of interference from emotional material in different kinds of contexts with variable temporal constrains.

Empathic reactions to the affective nature of the partners may in part explain our results, as the perception of a happy expression enhances subjective positive feelings (Van Kleef, 2009) and these increase cooperation tendencies (see for example, Kopelman et al., 2007), and the opposite may happen for anger (Forgas, 1998; Kopelman et al., 2006). Note, however, that empathic reactions in our experiments had to be context-dependent, as emotions did not influence explicit choices in Experiments 2 and 3, in which the behavior of participants was fully certain (see for example Frith & Frith, 2007; Ikezawa, Corbera & Wexler, 2013; Kadosh, Henson, Kadosh, Johnson & Dick, 2010). In addition, it could also be argued that priming mechanisms may have driven the obtained results. The mere presentation of happy facial expressions could promote cooperative (positive) behavior while anger could lead participants to the opposite (negative) reaction, regardless of the inferences or expectations of cooperative or non-cooperative action tendencies. To rule out such possibility it would be helpful to carry further experiments in which participants are warned that facial expressions are chosen at random and have nothing to do with the game partners. Although similar manipulations in previous studies ruled out the priming explanation (Gaertig et al., 2012), this effect may be relevant for the current paradigm.

Despite the conflict effect observed on RTs was present throughout the three experiments, it showed larger variability than it may be expected. Such variability could be reflecting individual differences in the resolution of emotional interference during interpersonal interactions. Bearing in mind that the temporal dynamics of responses were modeled based on each participant's pattern of response speed (i.e. bins), our results include a richer description of individual behavioral patterns. Indeed, the activation-suppression model (Ridderinkhof, 2002) has understanding individual variability by employing a distribution analysis as one of its goals (Forstmann, van den Wildenberg & Ridderinkhof, 2008). In addition, in the current paradigm conflict is measured by an interaction between two types of cooperation patterns and two emotional expressions, happiness and anger, which share some characteristics (Han et al., 2007; Lerner & Keltner, 2000), but also differ in many others (see for example, Todorov et al., 2008; Van Kleef et al., 2004). These differences could add variability to the data, which together with individual differences could lead to a more variable pattern of results. In any case, there is a consistent pattern of results on the core finding of the paper, that is, the interaction between the cooperation tendencies of the partners and their non-predictive emotional expression, which strongly suggests that ignoring emotional expressions is not sufficient to override their effect on the speed of cooperation responses.

The current experiments contain limitations that, although do not invalidate the results, should be tackled in future studies. In the first place, the game was played in a somehow artificial and iterative setting, which does not fully correspond to social interactions in daily life (e.g. Schilbach et al., 2013). This, however, was driven by the purpose of exploring the mandatory processing of salient emotions in an interference paradigm similar to those employed in the field of selective attention to study the automaticity of stimulus processing (see for example, Driver, 2001). The experimental approach allowed the use of faces previously rated in terms of the appropriateness of their emotional expressions (Tottenham et al., 2009), and also higher experimental control to measure subtle differences in the speed of responses. On the other hand, people nowadays engage in frequent social interactions that are artificial but still social in its nature (e.g. Facebook, WhatsApp; see Fischer & Reuber, 2011; Van Cleemput, 2010). In addition, humans display a natural propensity to interpret and represent stimuli in relation to their social content (e.g. Castelli, Happé, Frith & Frith, 2000; Sebanz, Knoblich & Prinz, 2003). To stress this, our cover story was conceived to match real social settings, in which the behavior of individuals displays different levels of trustworthiness. Furthermore, the artificial features of the design may have in fact reduced rather than increased the impact of facial expressions of emotions. These acquire full meaning in social contexts and thus it is reasonable to argue that their impact should be largest in live face-to-face situations (Barrett, 2012; Keltner & Gross, 1999; Parkinson et al., 2004; Parrott & Schulkin, 1993). In the second place, certain features of our task depart from the classic Trust Game, such as the lack of real payment to participants. Instructions, on the other hand, aimed at stressing the cooperative and non-cooperative tendencies of the partners, and thus asked participants to maximize benefits with each of them. Of note, previous studies in similar lines of research show that including small payments and/or omitting the instruction of benefits do not change the pattern of results (e.g. Gaertig et al., 2012). Future studies could incorporate videos with dynamic facial expressions, and also extend the range of emotions tested and the nature of the social settings. It would be interesting to explore whether the interference from ignored emotions also takes place in more natural settings, as well as the contextual dependency of the effects. In addition, it would be worth studying the extent of mandatory interpersonal emotional processing in people with deficits in emotional and/or social processing (e.g. alexithymia).

4.6. Conclusions

Overall, our findings are in line with social-emotional theories about how social constructs and emotions lead to different expectations about the proximal behavior of others (Fischer & Manstead, 2008; Haidt, 2001; Keltner & Haidt, 1999; Oosterhof & Todorov, 2008). In agreement with evolutionary claims about the existence of a specific *module* for judging trustworthiness (see for example Oosterhof & Todorov, 2008), the non-volitional processing of facial features aids in the prediction of what people are going to do next and other important social outcomes (Oosterhof & Todorov, 2008). The inconsistency between personal predispositions of cooperation and emotional facial information leads to opposite expectations and increases demands on decision-making, which are reflected in slower response times. Thus, in the same manner in which reading words cannot be avoided when attending to hue color in the classic Stroop task (Stroop, 1975), our results show for the first time that explicitly ignored emotions influence responses in a mandatory manner during interpersonal interactions.

4.7. References

Augoustinos, M., Walker, I., & Donaghue, N. (2014). *Social cognition: An integrated introduction*. Sage.

Averbeck, B. B., & Duchaine, B. (2009). Integration of social and utilitarian factors in decision making. *Emotion*, *9*, 599-608. http://psycnet.apa.org/DOI/10.1037/a0016509

Barrett, L. F. (2012). Emotions are real. *Emotion*, *12*, 413-429. http://psycnet.apa.org/DOI/10.1037/a0027555

Barrett, L. F., & Bliss--Moreau, E. (2009). Affect as a psychological primitive. *Advances in experimental social psychology*, *41*, 167-218. DOI: 10.1016/S0065-2601(08)00404-8

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, *10*, 122-142. DOI: 10.1006/game.1995.1027

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Campellone, T. R., & Kring, A. M. (2013). Who do you trust? The impact of facial emotion and behaviour on decision making. *Cognition & emotion*, *27*, 603-620. DOI:10.1080/02699931.2012.726608

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage*, *12*, 314-325. DOI: 10.1006/nimg.2000.0612

Dawkins, R. & Krebs, J. R. (1978). Animal signals: information or manipulation? In: J. R. Krebs & N. B. Davies (Eds.), *Behavioural Ecology: an Evolutionary Approach* (pp. 282-309), Oxford: Blackwell Scientific Publications.

Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature neuroscience*, *8*, 1611-1618. DOI:10.1038/nn1575

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*, 53-78.

Eckel, C. C., & Wilson, R. K. (2003). The human face of game theory: Trust and reciprocity in sequential games. *Trust and reciprocity:*

Interdisciplinary lessons from experimental research, 245-274. http://ssrn.com/abstract=1843493

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences*, 12, 374-80. DOI:10.1016/j.tics.2008.07.001

Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2008). Dissociable neural systems resolve conflict from emotional vs. non emotional distracters. *Cerebral Cortex*, *18*, 1475– 1484. DOI: 10.1093/cercor/bhm179

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. New Jersey: Prentice Hall.

Fehr, E., & Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in cognitive sciences*, 11, 419-427. DOI: 10.1016/j.tics.2007.09.002

Fischer, A. H., & Manstead, A. S. R. (2008). Social functions of emotion. In M. Lewis, J.

M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 456-468).

Fischer, E., & Reuber, A. R. (2011). Social interaction via new social media: (How) can interactions on Twitter affect effectual thinking and behavior?. *Journal of business venturing*, *26*, 1-18. DOI: 10.1016/j.jbusvent.2010.09.002

Forgas, J. P. (1991). Affective influences on partner choice: role of mood in social decisions. *Journal of personality and social psychology*, *61*, 708. DOI:10.1037/0022-3514.61.5.708

Forgas, J. P. (1994). The role of emotion in social judgments: An introductory review and an affect infusion model (AIM). *European Journal of Social Psychology*, *24*, 1-24. DOI: 10.1002/ejsp.2420240102

Forgas, J. P. (1995). Mood and judgment: the affect infusion model (AIM). *Psychological bulletin*, *117*, 39. http://dx.doi.org/10.1037/0033-2909.117.1.39

Forgas, J. P. (1998). On feeling good and getting your way: mood effects on negotiator cognition and bargaining strategies. *Journal of personality and social psychology*, *74*, 565. DOI: 10.1037/0022-3514.74.3.565

Forstmann, B., van den Wildenberg, W., & Ridderinkhof, K. (2008). Neural mechanisms, temporal dynamics, and individual differences in interference control. *Journal of Cognitive Neuroscience*, *20*, 1854-1865. DOI:10.1162/jocn.2008.20122

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought--action repertoires. *Cognition & Emotion*, *19*, 313-332. DOI: 10.1080/02699930441000238

Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. Academic Press. Frijda, N. H. (1988). The laws of emotion. *American psychologist*, *43*, 349. DOI:10.1037/0003-066X.43.5.349

Frijda, N.H., & Mesquita, B. (1994). The social roles and functions of emotions. In H.R. Markus & S. Kitayama (Eds.), *Emotion and Culture* (pp.51-87). New York: American Psychological Association.

Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology*, *17*, R724-R732. DOI: 10.1016/j.cub.2007.05.068

Funes, M. J., Lupiáñez, J., & Humphreys, G. (2010). Analyzing the generality of conflict adaptation effects. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 147. http://dx.doi.org/10.1037/a0017598

Gaertig, C., Moser, A., Alguacil, S., & Ruz, M. (2012). Social information and economic decision-making in the ultimatum game. *Frontiers in neuroscience*, 6:103. DOI: 10.3389/fnins.2012.00103

Gross, J. J. (1999). Emotion and emotion regulation. In L.A. Pervin & O.P. John (Eds.), *Handbook of personality: Theory and research* (pp. 525-552). New York: Guildford.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review*, *108*, 814-834. DOI: 10.1037/0033-295X.108.4.814

Han, S., Lerner, J. S., & Keltner, D. (2007). Feelings and consumer decision making: The appraisal-tendency framework. *Journal of consumer psychology*, *17*, 158-168. DOI:10.1016/S1057-7408(07)70023-2

Ikezawa, S., Corbera, S., & Wexler, B. E. (2013). Emotion self-regulation and empathy depend upon longer stimulus exposure. *Social cognitive and affective neuroscience*, nst148. DOI: 10.1093/scan/nst148

Kadosh, K. C., Henson, R. N., Kadosh, R. C., Johnson, M. H., & Dick, F. (2010). Task-dependent activation of face-sensitive cortex: an fMRI adaptation study. *Journal of Cognitive Neuroscience*, *22*, 903-917. DOI:10.1162/jocn.2009.21224

Keltner, D., & Gross, J. J. (1999). Functional accounts of emotions. *Cognition & Emotion*, *13*, 467-480. DOI:10.1080/026999399379140

Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition & Emotion*, *13*, 505-521. DOI:10.1080/026999399379168

Kopelman, S., Gewurz, I., & Sacharin, V. The Power of Presence: Strategic Response to Displayed Emotions in Negotiations. (2007). Ross School of Business, 1061; IACM 20TH Annual Conference. http://dx.DOI.org/10.2139/ssrn.951697

Kopelman, S., Rosette, A. S., & Thompson, L. (2006). The three faces of Eve: Strategic displays of positive, negative, and neutral emotions in negotiations. *Organizational Behavior and Human Decision Processes*, *99*, 81-101. DOI: 10.1016/j.obhdp.2005.08.003

Krebs, J. R., & Davies, N. B. (2009). *Behavioural ecology: an evolutionary approach*. John Wiley & Sons.

LeDoux, J. E. (2001). Emotion circuits in the brain. *The Science of Mental Health: Fear and anxiety*, 259.

Lee, V. K., & Harris, L. T. (2013). How social cognition can inform social decision making. *Frontiers in neuroscience*, 7:259. DOI: 10.3389/fnins.2013.00259

Lelieveld, G. J., Van Dijk, E., Van Beest, I., Steinel, W., & Van Kleef, G. A. (2011). Disappointed in you, angry about your offer: Distinct negative emotions induce concessions via different mechanisms. *Journal of Experimental Social Psychology*, 47, 635-641. DOI:10.1016/j.jesp.2010.12.015

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & Emotion*,*14*, 473-493. DOI**:** 10.1080/026999300402763

Lissek, S., Peters, S., Fuchs, N., Witthaus, H., Nicolas, V., Tegenthoff, M., ... & Brune, M. (2008). Cooperation and deception recruit different subsets of the theory-of-mind network. *PloS ONE*, 3, e2023. DOI: 10.1371/journal.pone.0002023

Loose, R., Kaufmann, C., Tucha, O., Auer, D. P., & Lange, K. W. (2006). Neural networks of response shifting: influence of task speed and stimulus material. *Brain research*, 1090, 146-155. DOI:10.1016/j.brainres.2006.03.039

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, 98, 11832-11835. DOI: 10.1073/pnas.211415698

Moser, A., Gäertig, C. & Ruz, M. (2014) Social information and personal interests modulate neural activity during economic decision-making. *Frontiers in Human Neuroscience*, 8:31. DOI: 10.3389/fnhum.2014.00031.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105, 11087-11092. DOI: 10.1073/pnas.0805664105

Parkinson, B., Fischer, A. H., & Manstead, A. S. (2004). *Emotion in social relations: Cultural, group, and interpersonal processes*. Psychology Press.

Parrott, W. G., & Schulkin, J. (1993). Neuropsychology and the cognitive nature of the emotions. *Cognition & Emotion*, 7, 43-59. DOI:10.1080/02699939308409175

Posner, M. I. (Ed.). (2011). *Cognitive neuroscience of attention*. Guilford Press.

Pourtois, G., Schettino, A., & Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological psychology*, 92, 492-512. DOI: 10.1016/j.biopsycho.2012.02.007

Reeck, C., & Egner, T. (2011). Affective privilege: asymmetric interference by emotional distracters. *Frontiers in psychology*, 2:232. DOI: 10.3389/fpsyg.2011.00232

Ridderinkhof, R. K. (2002). Micro-and macro-adjustments of task set: Activation and suppression in conflict tasks. *Psychological research*, 66, 312-323. DOI: 10.1007/s00426-002-0104-7

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Wijnen, J., & Burle, B. (2004). Response inhibition in conflict tasks is revealed in delta plots. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 369–377). New York: Guilford Press.

Ruz, M., Madrid, E., Tudela, P. (2013) Interactions between perceived emotions and executive attention in an interpersonal game. *Social, Cognitive & Affective Neuroscience*, 8, 838-44. DOI: 10.1093/scan/nss080

Ruz, M. & Moser, A. & Webster, K. (2011) Social expectations bias decision-making in uncertain interpersonal situations. *PLoS ONE* 6: e15762. DOI:10.1371/journal.pone.0015762.

Ruz, M. & Tudela, P. (2011) Emotional conflict in interpersonal interactions. *NeuroImage,* 54, 1685-91. DOI: 10.1016/j.neuroimage.2010.08.039

Sanfey, A. G. (2009). Expectations and social decision-making: biasing effects of prior knowledge on Ultimatum responses. *Mind & Society*, 8, 93-107. DOI: 10.1007/s11299-009-0053-6

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science,* 300, 1755-1758. DOI: 10.1126/science.1082976

Satpute, A. B., Badre, D., & Ochsner, K. N. (2013). Distinct regions of prefrontal cortex are associated with the controlled retrieval and selection of social information. *Cerebral Cortex*, bhs408. DOI: 10.1093/cercor/bhs408

Scharlemann, J. P., Eckel, C. C., Kacelnik, A., & Wilson, R. K. (2001). The value of a smile: Game theory with a human face. *Journal of*

Economic Psychology, 22, 617-640. DOI: 10.1016/S0167-4870(01)00059-9

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley,

K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 36, 393-414. DOI: 10.1017/S0140525X12000660

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime, Version 1.1. [Software]. Pittsburgh, PA: Psychology Software Tools.* Available from http://www.pstnet.com/support/kb.asp?TopicID=1508

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own?. *Cognition*, 88, B11-B21. DOI: 10.1016/S0010-0277(03)00043-X

Squire, R. F., Noudoost, B., Schafer, R. J., & Moore, T. (2013). Prefrontal Contributions to Visual Selective Attention. *Annual Review of Neuroscience*, 36, 451-466. DOI: 10.1146/annurev-neuro-062111-150439

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662. DOI:10.1037/h0054651

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: a model based approach. *Social cognitive and affective neuroscience*, 3, 119-127. DOI: 10.1093/scan/nsn009

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, 308, 1623-1626. DOI: 10.1126/science.1110589

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in cognitive sciences*, 12, 455-460. DOI: 10.1016/j.tics.2008.10.001

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and sociobiology*, 11, 375-424. DOI: 10.1016/0162-3095(90)90017-Z

Tortosa, M. I., Lupiáñez, J., & Ruz, M. (2013). Race, emotion and trust: An ERP study. *Brain research*, 1494, 44-55. DOI: 10.1016/j.brainres.2012.11.037

Tortosa, M. I., Strizhko, T., Capizzi, M., & Ruz, M. (2013). Interpersonal effects of emotion in a multi-round Trust Game. *Psicológica*, 34, 179-198. http://www.uv.es/psicologica/articulos2.13/3Tortosa.pdf

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., ... & Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry research*, 168, 242-249. DOI: 10.1016/j.psychres.2008.05.006

Van Cleemput, K. (2010). "I'll See You on IM, Text, or Call You": A Social Network Approach of Adolescents' Use of Communication Media. Bulletin of Science, *Technology & Society*, 30, 75-85. doi: 10.1177/0270467610363143

Van Kleef, G. A. (2009). How emotions regulate social life the emotions as social information (EASI) model. *Current Directions in Psychological Science*, 18, 184-188. doi: 10.1111/j.1467-8721.2009.01633.x

Van Kleef, G. A., De Dreu, C. K., & Manstead, A. S. (2004). The interpersonal effects of anger and happiness in negotiations. *Journal of personality and social psychology*, 86, 57. http://dx.doi.org/10.1037/0022-3514.86.1.57

Van Veen, V., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: a functional MRI study. *Neuroimage*, 27, 497-504. DOI:10.1016/j.neuroimage.2005.04.042

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in cognitive sciences*, 9, 585-594. DOI: 10.1016/j.tics.2005.10.011

Wegner, D. M., Schneider, D. J., Carter, S., & White, T. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology*, 53, 5-13. doi:10.1037/0022-3514.53.1.5

Capítulo 5: Serie Experimental III. Emotional conflict between personal identity and emotional expression: an fMRI study

El contenido de este capítulo está en preparación como Alguacil, S., Díaz, P., Kotz, S. A., Mestres-Missé, A., Tudela, P., & Ruz, M. Emotional conflict between personal identity and emotional expression: an fMRI study.

Abstract

Personal identity and facial emotional expressions serve as cues to make assumptions about the most likely behavior of other people during interpersonal interactions. These sources can also interfere each other when they lead to opposite predictions. We explored the neural mechanisms underlying conflict generated by ignored emotions when participants used identity to generate expectations about likely outcomes. Participants played a modified Trust Game where eight different facial identities were associated with gain or loss as probable outcomes. Faces displayed expressions of happiness or anger, which were non-predictive of future outcomes. Behavioral results showed the existence of conflict originating from the ignored emotional expressions. fMRI data showed that identities linked to gain strongly activated the dorsal striatum, involved in the formation of stimulus–action–reward associations. Congruent pairings activated the fusiform gyrus and other extrastriate areas relevant for face detection whereas incongruent interactions engaged the anterior cingulate cortex and the superior frontal gyrus, The response of the anterior cingulate cortex was functionally coupled with frontal and posteromedial regions, which may be related to the selection of appropriate actions trough the suppression of undesirable automatic responses on the basis of the evaluation of their motivational incentive value held in memory. Overall, our results highlight

the unavoidable influence of emotional information on identity-guided choices and show the involvement of brain regions linked to controlled processing and reward-based decision-making.

5.1. Introduction

Generating appropriate behaviors in a multifaceted environment requires multiple attentional mechanisms, which can be deployed in a voluntary or involuntary way. Contexts and most stimuli contain many different features, which sustain variable relevance depending on the task at hand. Those with which we have extensive experience are often associated with specific responses or action tendencies, which help to adjust behavior to the situation in a ballistic and efficient way (Norman & Shallice, 1986; Posner & Fan, 2004). However, if our goals target less practiced behaviors, the link between stimuli and responses is more fallible, and previously overlearned associations compete for behavioral control (Norman & Shallice, 1986; Posner & Fan, 2004). When different stimulus dimensions lead to opposite action tendencies, interference can generate the commission of errors (Kornblum, Hasbroucq, & Osman, 1990). To solve this interference, attention has to focus on relevant information while avoiding the influence of non-relevant elements. Thus, by implementing control mechanisms we can limit errors by selecting the most adequate response according to the situation.

In everyday life, the faces of other people are among the most relevant and frequent sources of information for humans, given our social nature. Facial information communicates internal states, attitudes and intentions, which are essential to forecast the actions of others and adapt our behavior accordingly (Ekman & Friesen, 2015; Fridlund, 2014; Frijda, 1988; Frijda & Mesquita, 1994; Frith & Frith, 2007). This information is extracted in an efficient manner from two different aspects embedded in faces: invariant and changeable features. Each of these, however, may lead to different mental expectations about what others are going to do the next, which generate ambiguous predictions as to which is the most appropriate behavior (Alguacil, Tudela, & Ruz, 2015). Focusing attention on the feature that holds the highest predictive power in the context, while ignoring the irrelevant surrounding information, could help elaborate the most accurate prediction.

Face processing is performed by a specialized distributed neural system composed by a Core and an Extended systems (Haxby & Gobbini, 2011; Haxby, Hoffman, & Gobbini, 2000). Each core region, under a hierarchical organization, sends decoded information as input to subsequent neural components in the system, with the goal of processing the visual appearance of faces. Shortly after face onset, the inferior occipital gyrus (occipital face area, OFA) processes its parts, and sends this information to both the lateral fusiform gyrus (FG) and superior temporal sulcus (STS). The FG decodes invariant features (i.e. identity), whereas the STS preferentially processes changeable features from faces (Winston, Henson, Fine-Goulden, & Dolan, 2004). Regions from the Extended System participate in representing personal knowledge associated to identity, and include the anterior paracingulate cortex, the temporoparietal junction, the anterior temporal cortex, the precuneus and the posterior cingulate cortex (Gobbini & Haxby, 2007). Areas such as the insula, the amygdala and the striatum are also involved in decoding the emotional content from facial information (Gobbini & Haxby, 2007; Haxby et al., 2000).

Despite the efforts made to explore emotional conflict occurrence and its resolution, there is little evidence of conflict stemming from faces embedded in social contexts. Most of the studies have employed two different modalities of information (Greening, Osuch, Williamson, & Mitchell, 2013; Kotz, Dengler, & Wittfoth, 2015; Watanabe et al., 2014; Watson et al., 2013) or have explored emotional interference by using classic stimulus-response paradigms (e.g. face-word Stroop task; Stroop, 1992) lacking social predictions of future outcomes (Deng et al., 2014; Egner, Etkin, Gale, & Hirsch, 2008; Etkin, Egner, Peraza, Kandel, & Hirsch, 2006; Gelder, Morris, & Dolan, 2005; Ochsner, Hughes, Robertson, Cooper, & Gabrieli, 2009; Torres-Quesada, Korb, Funes, Lupiáñez, & Egner, 2014). In the present study we evaluated the neural resources engaged by conflict between expectations generated from personal identities and their displayed emotions, when the former were predictors of positive and negative outcomes and the latter were non-predictive and thus irrelevant for the choices.

Although models of face processing suggest that identity and emotional expression are processed by independent neural routes, which would limit their interactions (Bruce & Young, 1986; Haxby et al., 2000), or would lead to asymmetric influences (Baudouin, Gilibert, Sansone, & Tiberghien, 2000; Kaufmann & Schweinberger, 2004; Posamentier & Abdi, 2003), many studies report data that suggest otherwise. For example, in the social realm, the overgeneralization hypothesis proposes that trustworthiness judgments about neutral faces are an extension of functionally adaptive systems for understanding the communicative meaning of emotional expressions that lead to behavioral approach and avoidance tendencies (Todorov, Mandisodza, Goren, & Hall, 2005; Todorov, Said, Engell, & Oosterhof, 2008).

According to the conflict-monitoring hypothesis (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Cohen, Botvinick, & Carter, 2000; MacDonald, Cohen, Stenger, & Carter, 2000), the ACC is involved in conflict detection. Its response to conflict serves to subsequent control regulation, triggering compensatory adjustments by the functional coupling of the ACC and dorsolateral prefrontal cortex (dlPFC; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Botvinick, Cohen, & Carter, 2004; Egner & Hirsch, 2005). After conflict detection, the dlPFC amplifies its activity to enhance task-relevant representations (Egner & Hirsch, 2005). Within this framework, theoretical models focused on the neural circuitry related to emotional conflict propose that its resolution comes from a top-down inhibition of amygdalar activity exerted by the rostral portion of ACC (Bush, Luu, & Posner, 2000; Etkin et al., 2006; Mohanty et al., 2007). Related investigations have explored if the strategic use of facial emotional expressions to guide economic decisions suffers from conflict when the natural action tendencies associated to emotions cue behaviors linked to the opposite expectations (Ruz & Tudela, 2011). This interference engaged the ACC, which increased its coupling with the anterior ventromedial prefrontal cortex and the precuneus, areas previously linked to theory of mind processes (Ruz & Tudela, 2011). More recently, some behavioral studies have shown that even when participants are explicitly asked to ignore the emotional expressions, as they carry no predictive power, their natural associations are triggered in an automatic way during alleged interpersonal interactions (Alguacil et al., 2015; Tortosa, Strizhko, Capizzi, & Ruz, 2013).

Therefore, the evidence so far supports the existence of a close relationship between the construction of inferences from facial identity and emotional expression, thus contradicting hypotheses about their limited or asymmetric mutual influence (Baudouin et al., 2000; Bruce & Young, 1986; Haxby et al., 2000; Kaufmann & Schweinberger, 2004; Posamentier & Abdi, 2003). It also shows that emotional expressions can be automatically processed and are likely to influence the decoding of facial identity (Alguacil et al., 2015; Tortosa et al., 2013). However, we still do not know the neural mechanisms that operate to resolve the interference stemming from ignored emotional content while predictive judgments are driven by personal identities. To study this, we employed a modified version of the Trust Game (TG) where participants played with partners whose identity was associated to expectations of gain or loss, and also displayed facial expressions of happiness or anger. They were also instructed to ignore facial expressions as they were irrelevant for the task and were non-predictive of future outcomes. We also explored the pattern of functional interactions between areas involved in conflict detection and other face-sensitive regions.

5.2. Materials and methods

5.2.1 Participants

Twenty-four right-handed healthy volunteers (12 females; mean age: 25; range: 20-36) took part in the experiment. All participants reported normal or corrected-to-normal vision. They all signed a consent form approved by the local Ethics Committee and received course credits in exchange for their participation.

5.2.2. Stimuli and procedure

Participants played a multiple-round computer adaptation of the Trust Game with 8 different partners (Berg, Dickhaut, & McCabe, 1995; Ruz & Tudela, 2011), depicted by photographs presented on the screen. Instructions explained that in every trial they had to decide whether to bet or not 1 fictitious EUR, received at the beginning of each round. If they decided to bet the money, then it was multiplied by 5 and passed to the partner. On the contrary, if they decided to keep it, they did not risk their initial monetary sum in the game and added that amount. Four facial identities cued a probable gain (2.5 EUR for the participant) and the other four a probable loss (no gains for that trial). Participants were told about the gain or loss predictions associated to each facial identity, and were asked to use this information during the task. The predictions afforded by identity were valid in 70% of the trials, and invalid in the remaining 30%. Instructions also explained that people in the photographs expressed happiness or anger, but that this emotional content was not relevant to any aspect of the game and, therefore, they had to be ignored. To facilitate the measurement of emotional conflict on the speed of responses, participants were asked to respond to the facial identity as quickly as possible. At the end of each round in the game, an asterisk or a hash symbol (either purple or blue) provided feedback indicative of the economic trial outcome (gain or loss).

Sixteen facial stimuli (8 identities, 4 females and 4 males) displaying happy or angry (50%) emotional expressions from the NimStim set (Tottenham et al., 2009) were used as photographs. Emotional expressions were orthogonal to facial identity and outcome (i.e. gain or loss). The predictive value of facial identities (gain or loss), the association between hand and response and the feedback symbols (and their ink color) were fully counterbalanced across participants.

The task and the stimuli were created and displayed using E-Prime 2.0 Professional software (Schneider, Eschman, & Zuccolotto, 2002), and viewed by the participants in the screen mounted at the back of the scanner by means of a mirror placed on top of the image acquisition coil. The delays between faces and feedback, and between trials, were jittered to allow the deconvolution of target and feedback related signals. Each trial comprised the following events (see Fig. 1.A): A symbol of EUR ('€'; 1.9° on average) displayed for 200 ms at the center of the screen, followed by a fixation point (+; 0.57° on average) of 500 ms. Then target face was flashed during 500 ms (7.15° on average) and after this replaced by an interval displaying the fixation point for 4.750-7.250 ms (the interval varied randomly in steps of 500 ms; mean = 6.000 ms). Afterwards a feedback symbol (‘*’, 0.67°, or ‘#’, 0.57°, on average; displayed in blue or purple ink) appeared for 500 ms, and lastly a second interval with the same characteristics as the previous one ended the trial. Participants were instructed to respond by pressing either with their index or middle finger of their right hand on a custom-made MRI-compatible button box. On average, a trial lasted 13.700 ms. The whole task was composed of 192 trials (48 per condition), for an approximate duration of 44 minutes. Before entering the scanner, participants performed a training block of 40 trials to familiarize themselves with the procedure. During this practice, the predictive value (i.e. gain, loss) of facial identities was always valid, to increase learning about their respective gain-loss associations. At the end of the task, participants filled a short questionnaire (10 point-Likert scale ranging from 0 to 10) to corroborate that were able to accurately indicate the associations between the facial identities and their association with gain and loss predictions.

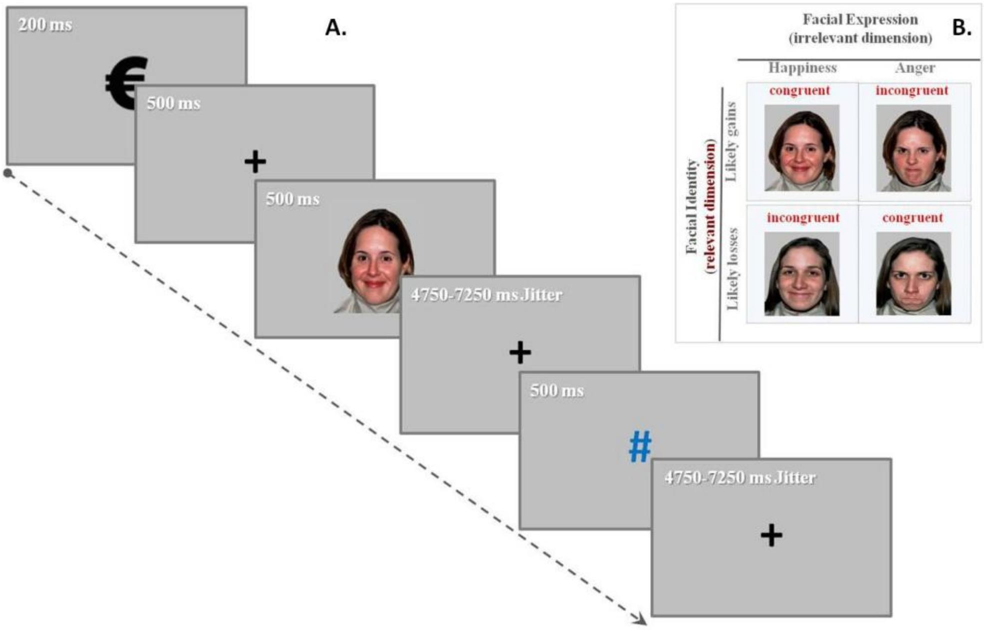

Figure 1. A. Sequence of events in a trial. **B.** The experimental design varied Prediction (of Gain or Loss, based on facial identities), Emotional Expression (Happy or Angry) and Prediction-Expression congruency (Congruent or Incongruent).

5.2.3. Image acquisition

Magnetic-resonance images were acquired using a Siemens Magnetom TrioTim 3T scanner at The Brain, Mind and Behavior Research Center (Spain). Functional images were obtained with a one-shot T2*-weighted echo planar imaging (EPI) sequence (time until echo [TE] = 30 ms, flip angle = 80 degrees, repetition time [TR] = 2 sec). Thirty-two descendent sagittal slices with a thickness of 3.5 mm (20% gap) covered the entire brain (64 x 64 matrix with a field of view of 224 mm, voxel size of 3.5 x 3.5 x 3.5

mm). The event-related experiment was performed in a single run consisting of 1325 volumes. The first 5 images were discarded to allow for saturation of the signal. In addition, we acquired a standard structural image of each participant using a high-resolution T1-weighted sequence (TR = 1900 ms; TE = 2.52 ms; 1 x 1 mm in-plane resolution and 1 mm slice thickness).

5.2.4. Analyses

Image preprocessing and analysis was performed with SPM8 (Welcome Department of Imaging Neuroscience, University of London, UK; http://www.fil.ion.ucl.ac.uk/spm). First, images were realigned and unwarped using a least-squares approach and a six-parameter (rigid body) spatial transformation to correct for motion artifacts. High-resolution anatomical T1 images were then coregistered with the realigned functional images (Friston, Frith, Frackowiak, & Turner, 1995). Functional images were spatially normalized to the standard EPI template included in SPM8, and smoothed using an 8 mm full-width at half-maximum isotropic Gaussian kernel. A 128 sec high-pass filter was used to remove low-frequency artifacts.

First-level statistical analysis was performed with a General Linear Model for each participant with corrections for serial autocorrelations using the AR(1) model. It included eight regressors for the target face and feedback symbol (four for each one), both corresponding to predictions of gain associated to happy or angry expressions (GH and GA), predictions of loss linked to happy or angry expressions (LH and LA), valid and invalid feedback of gain (FV and FI) and valid and invalid feedback of loss (FI and FI). These regressors were convolved with the standard hemodynamic response function and modeled as events with zero duration. Contrasts of interest for the face target (one for every regressor in the model: GH, GA, LH, LA) were obtained for each participant and then entered into a 2 x 2 (Prediction: Gain vs. Loss; Emotional Expression: Happy vs. Angry) flexible factorial design at the second-level analysis. Only clusters surviving a family-wise error (FWE) correction thresholded at a 0.05 cluster-level (initial cluster-forming threshold was uncorrected $p<0.001$) are reported. The flexible factorial analysis used a "subject by condition" design to model the interactions between participants and conditions (GH; GA; LH; LA) and also the interaction between Prediction and Emotional Expression. The contrasts aimed to identify the brain areas involved in the processing of facial identities predicting gain or loss (i.e. G > L and vice versa), and also related to the processing of the emotional content of the identities (i.e. H > A and vice versa). Besides, we employed the interaction between Prediction and Emotional Expressions (see Fig. 1.B) to study the brain regions involved in conflict between congruent (GH/LA > GA/LH) and incongruent Prediction/Emotional Expression pairings (GA/LH > GH/LA).

We also performed a Psychophysiological interactions (Friston et al., 1997) analysis to assess the networks involved in conflict processing . The flexible factorial second-level analysis previously described yielded a significant activation in the Anterior Cingulate Cortex (ACC; x = 3, y = 17, z = 19) for the incongruent Prediction/Emotional Expression pairings (GA/LH > GH/LA). This coordinate was the center of 6-mm radius spherical seeds in the subsequent PPI analysis. A first-level GLM analysis was then performed for each of the four variables of interest (GH, GA, LH, LA), using the ACC time course, the psychological variables (Prediction and Emotional Expression) and the interaction term (PPI) obtained in SPM8 as regressors. After obtaining a contrast image of the interaction (GA/LH > GH/LA) for each subject, a second-level 2 x 2 flexible factorial analysis was performed (Prediction: Gain vs. Loss; Emotional Expression: Happy vs. Angry). Clusters surviving the statistical threshold corresponded to patterns of activations that covaried with the activity in the ACC during conflict trials (GA/LH > GH/LA).

We used MRIcroGL software (Rorden & Brett, 2000) to display the activations in the figures.

5.3. Results

5.3.1. Behavioral

Only trials in which participants did not respond were discarded from the analyses (0.45 %). On average, participants betted their money on 48 % of the trials. The ANOVA on the mean betting rates yielded a main effect of Prediction, F (1, 23) = 700.71, p < .001, ηp^2 = .97, as participants betted their money a higher number of times when they were expecting Gain, M = 89 %, SE = 2.3 %, compared to Loss, M = *55* %, SE = 4.5 %. The ANOVA also revealed an effect of Emotional Expression close to significance levels, F (1, 23) = 3.883, p = .061, ηp^2 = .14, as participants betted more money when they were watching a face with an expression of happiness, M = 49 %, SE = 1.1 %, than when it was displaying anger, M = 47 %, SE = 1.2 %. The interaction between Prediction and Emotional Expression did not reach significance, F (1, 23) = 1.117, p = .301.

Reaction times (RTs) analysis revealed an interaction effect between the Prediction and the Emotional Expression factors, F (1, 23) = 13.83, p = .001, ηp^2 = .37, revealing a clear effect of emotional conflict. Identities predicting gain and displaying happiness generated faster choices, M = 809.61, SE = 34.68, than those expressing anger, M = 838.971, SE = 31.53, F (1, 23) = 5.24, p < .05, ηp^2 = .19. In an opposite fashion, identities predicting loss received faster responses when they expressed anger, M = 813.85, SE = 32.56, than when they displayed happiness, M = 842.25, SE =

34.14, F (1, 23) = 10.71, p < .01, ηp^2 = .32 (see Fig. 2). There were no other main effects or interactions, all ps > .848.

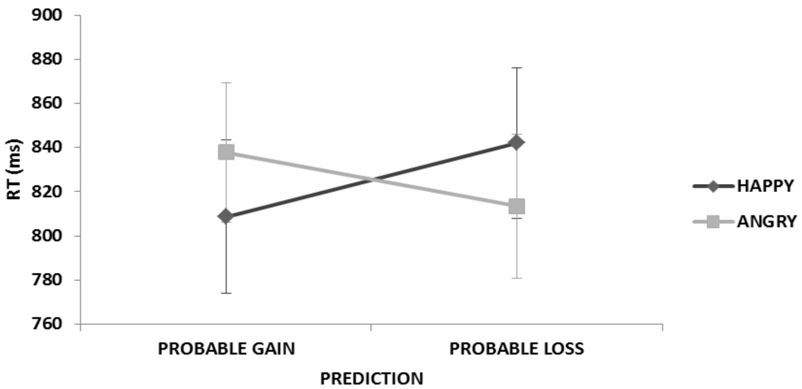

Figure 2. Conflict effect as shown by RTs. Error bars represent the standard error of the mean.

5.3.2. Neuroimaging

5.3.2.1. Flexible Factorial

Identities predicting gain (G > L) generated activity in a large cluster (k = 921) including the right (15, -7, 22) and left body of the caudate (rBC/lBC; -15, -4, 22) and left putamen (lPUT; -18, 17, -5). No significant differences were found for the opposite contrast, L > G, or Emotional Expressions, at the specified threshold (see Fig. 3).

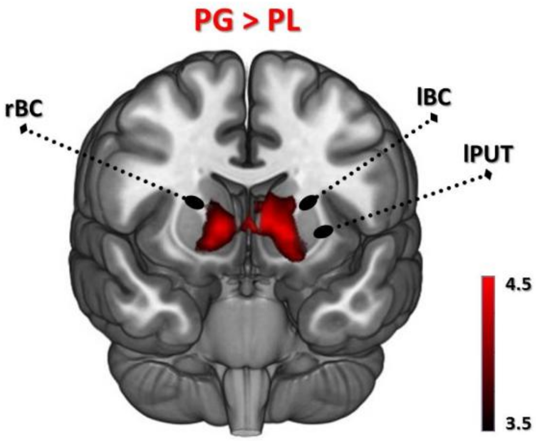

Figure 3: Coronal display of activations obtained for the G > L contrast. The scale represents peak *t* values.

Trials in which the prediction and emotional expressions were congruent (GH/LA > GA/LH; see Fig. 4.A) activated the right fusiform gyrus (rFG; 33, -58, -5), the left lingual gyrus (lLING; -21, -76, -2) and also the left cuneus (lCUN; -12, -97, -5), *k* = 440. Incongruent situations (GA/LH > GH/LA; see Fig. 4.B), on the other hand, yielded significant activity in a cluster (*k* = 98) located at the ACC (3, 17, 19). In addition, a second cluster (*k* = 101) included the left dACC (-12, 35, 22) and the left medial superior frontal gyrus (lSFGmed; -12, 7, 52).

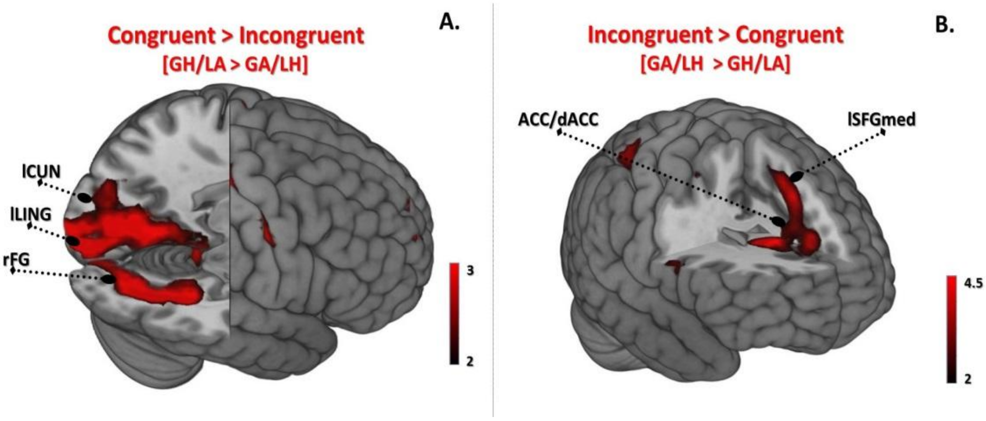

Figure 4: Emotional conflict analysis results for congruent (A) and incongruent (B) trials. Scales represent peak *t* values.

5.3.2.2. Connectivity

Incongruent situations increased the connectivity between the ACC with a cluster (*k* = 111) containing the left dorsal posterior cingulate cortex (dPCC, -6, -64, 25) and the left precuneus (lPCUN; -12, -58, 31). Conflict between facial cues also enhanced the connectivity of the ACC with a cluster (*k* = 133) including the left inferior frontal gyrus on its orbital (LORBinf; -33, 32, -8) and triangular (LINFtriang; -48, 26, 4) regions (see Fig. 5).

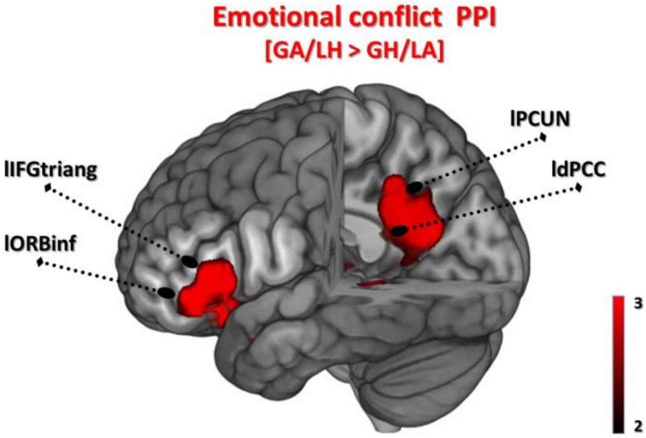

Figure 5: Results of the psychophysiological interactions analysis (PPI) seeded in the anterior cingulate cortex (ACC; x = 3, y = 17, z = 19) displaying increased connectivity during incongruent trials. The scale represents peak *t* values.

5.4. Discussion

In the present investigation we studied the neural responses to conflict in a decision-making scenario where participants used the identity of faces to predict future outcomes while ignoring their irrelevant facial expressions. Results showed that even though participants mostly based their choices on identity information, there was an effect of conflict arising from the expectations linked to the ignored emotional expressions. Decisions were slower when the two sources of information led to incongruent, non-overlapping, action tendencies. When participants faced an identity that anticipated probable gain, the body of the caudate and the putamen were activated. Congruent identity and emotion predictions engaged the cuneus, lingual and fusiform giri, whereas incongruent combinations increased activity in the ACC and SFG. In addition, the functional interactions of the ACC and IFG, precuneus and PCC increased during conflictive choices.

In our experiment, participants had to learn the associations between each facial identity and its particular value of probable reward/non-reward (i.e. probable gain/loss) and held them in memory. In relation to theories about reward-based learning and decision-making, such associations help individuals to choose the most advantageous option by comparing the values of predicted outcomes of the different actions toward stimuli (Balleine, Delgado, & Hikosaka, 2007; FitzGerald, Friston, & Dolan, 2012). Thus, the selection of actions partially depends on their probable consequences; actions linked to advantageous outcomes will be preferentially selected. In accord to this, behavioral data in our study showed that participants strategically betted with higher probability when the identity had a greater chance to lead to gain than when it anticipated a probable loss. fMRI results showed that identities signaling gain activated both the caudate body and the putamen, areas embedded in the dorsal striatum. These regions have been described as relevant to the formation of stimulus-action-reward associations and to motivational processes in goal-directed behavior (Balleine et al., 2007; FitzGerald et al., 2012; Haruno & Kawato, 2006; Peterson & Seger, 2013), although they may play dissociable roles in those processes (Haruno &

Kawato, 2006; Peterson & Seger, 2013). Our results add evidence in favor of the involvement of the dorsal striatum in successful learning of stimuli-specific value to anticipate the most advantageous outcome after training (Balleine et al., 2007; FitzGerald et al., 2012; Peterson & Seger, 2013). However, the opposite contrast (loss more than gain) did not reveal any effect. In social decision-making contexts, activity in the striatum has also been found increased during mutually cooperative social interactions compared to non-cooperation. Hence, the neural response of this region seems to be sensitive to the reward value of cooperative acts (Rilling et al., 2002). Thus, the lack of effect for identities associated with loss could be possibly explained by a higher attribution of cooperative value to identities linked to predictions of gain.

The interaction between predictive personal identities and their ignored emotional expressions led to several findings. When expectations from these two sources were congruent, leading to overlapping action tendencies, activity increased in the left cuneus and lingual gyrus, and also in the right FG. These three regions have been related to the detection of faces (mostly the lingual gyrus), and the recognition of invariant and changeable aspects (Lai & Wu, 2013; Nomi et al., 2008). The response of these three areas, jointly with activity in others such as the precuneus and the parahippocampal gyrus, has been observed during judgments where the stability of the spatial distribution of the information contained in the stimulus is important (Sulpizio, Committeri, Lambrey, Berthoz, & Galati, 2013). For example, Sulpizio and colleagues (2013) found significant activity in the lingual gyrus during the classification of stimuli with spatial configurations leading to two possible interpretations; the neural response of this area increased during the presentation periods in which the configuration remained stable. During congruent situations, this region might be important to focus attention on stable configural facial features. In our study, identity was a stable (i.e. invariant) feature that helped participants to choose the more appropriate option. Thus, on congruent contexts, participants could easily focus on this information by attending to the identity without emotional interference. On the other hand, the cuneus has been related to the regulation of autonomic responses and to the early perception of emotional information (Lai & Wu, 2013; O'Connor, Gündel, McRae, & Lane, 2007). It could be that this region contributes to the regulation of the processing of emotional information when it is congruent with expectations emerged from facial identity during early visual inspection, regardless of its specific valence. Finally, the role of the right FG on facial identity decoding has been extensively probed (Gobbini & Haxby, 2007; Haxby & Gobbini, 2011; Haxby et al., 2000; Haxby, Hoffman, & Gobbini, 2002). In the absence of conflict between expectations from invariant and changeable facial aspects, the selection of the most advantageous option would be supported by detecting and attending to the relevant invariant features (i.e. identity). This could help to choose among the facial identities held in memory and to generate the most appropriate response taking into consideration the expectations linked to the detected identity. These results suggest that a preliminary analysis of emotional salient stimuli affects activity in perceptual regions linked to facial processing, increasing their response when different facial aspects leads to action tendencies of the same valence.

Conversely, when identity-based expectations conflicted with the natural tendencies associated to the ignored emotions, activations increased in conflict and control-related areas such as the ACC, including its dorsal portion, and the left IFG. As discussed earlier, the ability to detect the occurrence of conflict in information processing, for example response competition, has been mainly attributed to the ACC, and more specifically to its dorsal portion for domain-general conflict and global interference (Botvinick et al., 1999; Etkin et al., 2006; Ochsner et al., 2009; Rushworth, Walton, Kennerley, & Bannerman, 2004; Torres-Quesada et al., 2014; Zaki, Hennigan, Weber, & Ochsner, 2010). During monitoring of conflict, the ACC triggers a signal that calls for strategic adjustments in control. These serve to reduce conflict in subsequent performance (Botvinick et al., 2004, 1999). The ACC has also been shown to be relevant for the evaluation of outcomes following a specific action, by encoding the connection between actions and the value, helpful to guide decisions (Botvinick et al., 1999; Rushworth & Behrens, 2008; Rushworth et al., 2004). Besides the ACC activation, incongruent situations also increased activity in the medial SFG, a region that has been related to the selection of action sets (Rushworth et al., 2004). Its activity may reflect the orienting of attention to the relevant target location (Botvinick et al., 1999; Casey et al., 2000), and is also greater for incongruent situations (Milham et al., 2001). On this line, the spatial interrelation between facial parts has been extensively probed as a crucial element for identity recognition (Maurer, Grand, & Mondloch, 2002; Rotshtein, Geng, Driver, & Dolan, 2007). In our experiment, whereas emotions were non-predictive of outcomes, identity was predictive on a high percentage (70%) of trials. Thus, it is reasonable to assume that when the display of emotions hindered the processing of the relevant identity information, participants may have reoriented attention to the spatial distribution of the parts of the face to focus on identity. Hence it could be argued that, in our experiment, when expectations from facial features and emotions led to incongruent action tendencies, the conflict signal triggered by the ACC could initiate adjustment in control by reorienting attention, with participation of the SFG, to spatial locations relevant to select the most appropriated action.

The coactivation of ACC and the SFG was accompanied by a functional coupling of the former with the dorsal PCC, the left precuneus and the left IFG. Although most of the studies have stressed the role of the right IFG in control, its homologous region in the left hemisphere has also been shown as a critical area for suppression of prepotent but inappropriate responses (Menon, Adleman, White, Glover, & Reiss, 2001; Swick, Ashley, & Turken, 2008; Wittfoth, Küstermann, Fahle, & Herrmann, 2008). In coherence with this role, activity in the left IFG may reflect the need of selection among competing alternatives from memory (Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). The other two regions functionally interconnected with ACC were the PCC and the precuneus. These areas have been related to a wide range of social and cognitive conflict processing (Cavanna & Trimble, 2006; Pearson, Heilbronner, Barack, Hayden, & Platt, 2011). Both are also recruited during successful memory retrieval, mostly for episodic memories, and mental imagery strategies (Cavanna & Trimble, 2006; Maddock, Garrett, & Buonocore, 2001), They are relevant for social cognition and emotion processing in moral judgments, and play important evaluative functions in monitoring behavior and self-relevance (Wittfoth et al., 2008). In cooperation scenarios, the greater the degree of trustworthiness and mentalizing about social cues, the stronger the neural response of the PCC and the precuneus (Fett, Gromann, Giampietro, Shergill, &

Krabbendam, 2012). In this context, the PCC shows a strong neural response when moral judgments are difficult and entail some sort of conflict (Greene, Nystrom, Engell, Darley, & Cohen, 2004; Johnson et al., 2006). This region, a relevant node in the default network (e.g. Leech, Kamourieh, Beckmann, & Sharp, 2011), is supposed to respond strongly during initial learning and also during environmental changes that require a subsequent selection of a new set of behaviors (Pearson et al., 2011). This is coherent with the idea that the PCC together with the precuneus and prefrontal cortices is involved in the constant gathering of information and in the presentation of both internal and external world (Gusnard & Raichle, 2001). On this regard, the PCC has been proposed as a crucial neural connection between emotion and top-down control of attention as it is strongly responsive to those situations where motivational incentives exert a visual spatial bias (Small et al., 2005). Consistent with this approach, in contrast to the executive role attributed to the ACC, the PCC play a more evaluative function in the current task settings (Vogt, Finch, & Olson, 1992). Although speculative, we propose that the selection of specific action sets, to appropriately respond to the conflict, is made by comparing among the representations of each identity and the reward value of its outcome. As these representations are acquired during the instruction phase and are held in episodic memory, thus regions related to evaluation and episodic memory retrieval are involved (i.e. PCC and precuneus).

In short, decisions guided by facial identity are biased by the reward value (i.e. gain or loss) of action-outcomes linked to each identity. Facial identities predicting gain strongly activated relevant areas from the dorsal striatum involved in the formation of stimulus-action-reward associations. When their decoding of was not hindered by irrelevant and non-predictive emotional information, then occipital and temporal areas from the distributed neural system for face processing, relevant for the configuration of the identity, implemented the most adequate set of actions (Gobbini & Haxby, 2007; Haxby & Gobbini, 2011; Haxby et al., 2000, 2002). However, incongruent situations, call for additional conflict detection and control implementation mechanisms operated by a network including middle-frontal, frontolateral and a posteromedial areas. Our study, for very first time, shows how control mechanisms operate when emotional conflict emerges from incongruent expectations and non-overlapping action tendencies associated to invariant and changeable facial aspects during personal interactions.

Although we found emotional conflict effects, results did not reveal any difference in activations for happy and angry emotions. The lack of predictive value of these expressions could be a plausible explanation for this, although this type of null result is not uncommon in emotion research (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012), in agreement with constructivists theories of emotion processing (Barrett, 2013). Future studies could compare the current paradigm with situations in which emotions do carry predictive information regarding the behavior of partners in interacting situations. Also, although previous research suggests that offering money payment to participants does not change the pattern of results (Alguacil, Madrid, Tudela & Ruz, *in preparation*; Gaertig, Moser, Alguacil, & Ruz, 2012), future studies could include real economic incentives to performance. It would also be interesting to employ mixed designs to discriminate between sustained and phasic control mechanisms and explore the role of these networks in representing the two types of interpersonal cues and their conflict. Finally, it would be informative to create paradigms to understand better the specific role of different limbic areas during emotional conflict processing.

5.5. References

Alguacil, S., Tudela, P., & Ruz, M. (2015). Ignoring facial emotion expressions does not eliminate their influence on cooperation decisions.

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The Role of the Dorsal Striatum in Reward and Decision-Making. *The Journal of Neuroscience*, *27*(31), 8161–8165. http://doi.org/10.1523/JNEUROSCI.1554-07.2007

Barrett, L. F. (2013). Psychological Construction: The Darwinian Approach to the Science of Emotion. *Emotion Review*, *5*(4), 379–389. http://doi.org/10.1177/1754073913489753

Baudouin, J. Y., Gilibert, D., Sansone, S., & Tiberghien, G. (2000). When the smile is a cue to familiarity. *Memory (Hove, England)*, *8*(5), 285–292. http://doi.org/10.1080/09658210050117717

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, *10*(1), 122–142. http://doi.org/10.1006/game.1995.1027

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652.

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. http://doi.org/10.1016/j.tics.2004.10.003

Botvinick, M. M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*(6758), 179–181. http://doi.org/10.1038/46035

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305–327. http://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*(6), 215–222.

Casey, B. J., Thomas, K. M., Welsh, T. F., Badgaiyan, R. D., Eccard, C. H., Jennings, J. R., & Crone, E. A. (2000). Dissociation of response conflict, attentional selection, and expectancy with functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(15), 8728–8733.

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, *129*(3), 564–583. http://doi.org/10.1093/brain/awl004

Cohen, J. D., Botvinick, M. M., & Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: who's in control? *Nature Neuroscience*, *3*(5), 421–423. http://doi.org/10.1038/74783

Deng, Z., Wei, D., Xue, S., Du, X., Hitchman, G., & Qiu, J. (2014). Regional gray matter density associated with emotional conflict resolution: evidence from voxel-based morphometry. *Neuroscience*, *275*, 500–507. http://doi.org/10.1016/j.neuroscience.2014.06.040

Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2008). Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(6), 1475–1484. http://doi.org/10.1093/cercor/bhm179

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, *8*(12), 1784–1790. http://doi.org/10.1038/nn1594

Ekman, P., & Friesen, W. V. (2015). *Unmasking the Face: A Guide to Recognizing Emotions From Facial Expressions*. Malor Books.

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron*, *51*(6), 871–882. http://doi.org/10.1016/j.neuron.2006.07.029

Fett, A.-K. J., Gromann, P. M., Giampietro, V., Shergill, S. S., & Krabbendam, L. (2012). Default Distrust? An fMRI investigation of the neural development of trust and cooperation. *Social Cognitive and Affective Neuroscience*, nss144. http://doi.org/10.1093/scan/nss144

FitzGerald, T. H. B., Friston, K. J., & Dolan, R. J. (2012). Action-specific value signals in reward-related regions of the human brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(46), 16417–16423a. http://doi.org/10.1523/JNEUROSCI.3254-12.2012

Fridlund, A. J. (2014). *Human Facial Expression: An Evolutionary View*. Academic Press.

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*(5), 349–358. http://doi.org/10.1037/0003-066X.43.5.349

Frijda, N. H., & Mesquita, B. (1994). The social roles and functions of emotions. In S. Kitayama & H. R. Markus (Eds.), *Emotion and culture: Empirical studies of mutual influence* (pp. 51–87). Washington, DC, US: American Psychological Association.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, *6*(3), 218–229. http://doi.org/10.1006/nimg.1997.0291

Friston, K. J., Frith, C. D., Frackowiak, R. S., & Turner, R. (1995). Characterizing dynamic brain responses with fMRI: a multivariate approach. *NeuroImage*, *2*(2), 166–172.

Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology: CB*, *17*(16), R724–732. http://doi.org/10.1016/j.cub.2007.05.068

Gaertig, C., Moser, A., Alguacil, S., & Ruz, M. (2012). Social Information and Economic Decision-Making in the Ultimatum Game. *Frontiers in Neuroscience*, *6*. http://doi.org/10.3389/fnins.2012.00103

Gelder, B. de, Morris, J. S., & Dolan, R. J. (2005). Unconscious fear influences emotional awareness of faces and voices. *Proceedings of* *the National Academy of Sciences of the United States of America*, *102*(51), 18682–18687. http://doi.org/10.1073/pnas.0509179102

Gobbini, M. I., & Haxby, J. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, *45*(1), 32–41. http://doi.org/10.1016/j.neuropsychologia.2006.04.015

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The Neural Bases of Cognitive Conflict and Control in Moral Judgment. *Neuron*, *44*(2), 389–400. http://doi.org/10.1016/j.neuron.2004.09.027

Greening, S. G., Osuch, E. A., Williamson, P. C., & Mitchell, D. G. V. (2013). Emotion-related brain activity to conflicting socio-emotional cues in unmedicated depression. *Journal of Affective Disorders*, *150*(3), 1136–1141. http://doi.org/10.1016/j.jad.2013.05.053

Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nature Reviews. Neuroscience*, *2*(10), 685–694. http://doi.org/10.1038/35094500

Haruno, M., & Kawato, M. (2006). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *Journal* *of Neurophysiology*, *95*(2), 948–959. http://doi.org/10.1152/jn.00382.2005

Haxby, J., & Gobbini, M. I. (2011). Distributed Neural Systems for Face Perception. In A. Calder, G. Rhodes, M. Johnson, & J. Haxby (Eds.), *Oxford Handbook of Face Perception* (pp. 93–110). OUP Oxford.

Haxby, J., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. http://doi.org/10.1016/S1364-6613(00)01482-0

Haxby, J., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51*(1), 59–67.

Johnson, M. K., Raye, C. L., Mitchell, K. J., Touryan, S. R., Greene, E. J., & Nolen-Hoeksema, S. (2006). Dissociating medial frontal and posterior cingulate activity during self-reflection. *Social Cognitive and Affective Neuroscience*, *1*(1), 56–64. http://doi.org/10.1093/scan/nsl004

Kaufmann, J. M., & Schweinberger, S. R. (2004). Expression influences the recognition of familiar faces. *Perception*, *33*(4), 399–408.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility--a model and taxonomy. *Psychological Review*, *97*(2), 253–270.

Kotz, S. A., Dengler, R., & Wittfoth, M. (2015). Valence-specific conflict moderation in the dorso-medial PFC and the caudate head in emotional speech. *Social Cognitive and Affective Neuroscience*, *10*(2), 165–171. http://doi.org/10.1093/scan/nsu021

Lai, C.-H., & Wu, Y.-T. (2013). Decreased regional homogeneity in lingual gyrus, increased regional homogeneity in cuneus and correlations with panic symptom severity of first-episode, medication-naïve and late-onset panic disorder patients. *Psychiatry Research: Neuroimaging*, *211*(2), 127–131. http://doi.org/10.1016/j.pscychresns.2012.11.006

Leech, R., Kamourieh, S., Beckmann, C. F., & Sharp, D. J. (2011). Fractionating the default mode network: distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(9), 3217–3224. http://doi.org/10.1523/JNEUROSCI.5626-10.2011

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *The Behavioral and Brain Sciences*, *35*(3), 121–143. http://doi.org/10.1017/S0140525X11000446

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science (New York, N.Y.), 288*(5472), 1835–1838.

Maddock, R. J., Garrett, A. S., & Buonocore, M. H. (2001). Remembering familiar people: the posterior cingulate cortex and autobiographical memory retrieval. *Neuroscience*, *104*(3), 667–676.

Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260.

Menon, V., Adleman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related brain activation during a Go/NoGo response inhibition task. *Human Brain Mapping*, *12*(3), 131–143.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., & Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Brain Research. Cognitive Brain Research*, *12*(3), 467–473.

Mohanty, A., Engels, A. S., Herrington, J. D., Heller, W., Ho, M.-H. R., Banich, M. T., … Miller, G. A. (2007). Differential engagement of anterior cingulate cortex subdivisions for cognitive and emotional function. *Psychophysiology*, *44*(3), 343–351. http://doi.org/10.1111/j.1469-8986.2007.00515.x

Nomi, J. S., Scherfeld, D., Friederichs, S., Schäfer, R., Franz, M., Wittsack, H.-J., … Seitz, R. J. (2008). On the neural networks of empathy: A principal component analysis of an fMRI study. *Behavioral and Brain Functions*, *4*, 41. http://doi.org/10.1186/1744-9081-4-41

Norman, D., & Shallice, T. (1986). Attention to Action: Willed and Automatic Control of Behavior. In R. Davidson, R. Schwartz, & D. Shapiro (Eds.), *Consciousness and Self-Regulation: Advances in Research and Theory IV*. Plenum Press.

Ochsner, K. N., Hughes, B., Robertson, E. R., Cooper, J. C., & Gabrieli, J. D. E. (2009). Neural systems supporting the control of affective and cognitive conflicts. *Journal of Cognitive Neuroscience*, *21*(9), 1842–1855. http://doi.org/10.1162/jocn.2009.21129

O'Connor, M.-F., Gündel, H., McRae, K., & Lane, R. D. (2007). Baseline vagal tone predicts BOLD response during elicitation of grief. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *32*(10), 2184–2189. http://doi.org/10.1038/sj.npp.1301342

Pearson, J. M., Heilbronner, S. R., Barack, D. L., Hayden, B. Y., & Platt, M. L. (2011). Posterior cingulate cortex: adapting behavior to a changing world. *Trends in Cognitive Sciences*, *15*(4), 143–151. http://doi.org/10.1016/j.tics.2011.02.002

Peterson, E. J., & Seger, C. A. (2013). Many hats: intratrial and reward level-dependent BOLD activity in the striatum and premotor cortex. *Journal of Neurophysiology*, *110*(7), 1689–1702. http://doi.org/10.1152/jn.00164.2012

Posamentier, M. T., & Abdi, H. (2003). Processing Faces and Facial Expressions. *Neuropsychology Review*, *13*(3), 113–143. http://doi.org/10.1023/A:1025519712569

Posner, M. I., & Fan, J. (2004). Attention as an Organ System. In J. R. Pomerantz (Ed.), *Topics in Integrative Neuroscience: From Cells to Cognition* (Vol. 9780521869133, pp. 31–61). Carnbridge University Press.

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A Neural Basis for Social Cooperation. *Neuron*, *35*(2), 395–405. http://doi.org/10.1016/S0896-6273(02)00755-9

Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology*, *12*(4), 191–200.

Rotshtein, P., Geng, J. J., Driver, J., & Dolan, R. J. (2007). Role of Features and Second-order Spatial Relations in Face Discrimination, Face Recognition, and Individual Face Skills: Behavioral and Functional Magnetic Resonance Imaging Data. *Journal of Cognitive Neuroscience*, *19*(9), 1435–1452. http://doi.org/10.1162/jocn.2007.19.9.1435

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*(4), 389–397. http://doi.org/10.1038/nn2066

Rushworth, M. F. S., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences*, *8*(9), 410–417. http://doi.org/10.1016/j.tics.2004.07.009

Ruz, M., & Tudela, P. (2011). Emotional conflict in interpersonal interactions. *NeuroImage, 54*(2), 1685–1691. http://doi.org/10.1016/j.neuroimage.2010.08.039

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*. Psychology Software Incorporated.

Small, D. M., Gitelman, D., Simmons, K., Bloise, S. M., Parrish, T., & Mesulam, M.-M. (2005). Monetary incentives enhance processing in brain regions mediating top-down control of attention. *Cerebral Cortex (New York, N.Y.: 1991)*, *15*(12), 1855–1865. http://doi.org/10.1093/cercor/bhi063

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. http://doi.org/10.1037/h0054651

Sulpizio, V., Committeri, G., Lambrey, S., Berthoz, A., & Galati, G. (2013). Selective role of lingual/parahippocampal gyrus and retrosplenial complex in spatial memory across viewpoint changes relative to the environmental reference frame. *Behavioural Brain Research*, *242*, 62–75. http://doi.org/10.1016/j.bbr.2012.12.031

Swick, D., Ashley, V., & Turken, A. U. (2008). Left inferior frontal gyrus is critical for response inhibition. *BMC Neuroscience*, *9,* 102. http://doi.org/10.1186/1471-2202-9-102

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(26), 14792–14797.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science (New York, N.Y.),* *308*(5728), 1623–1626. http://doi.org/10.1126/science.1110589

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*(12), 455–460. http://doi.org/10.1016/j.tics.2008.10.001

Torres-Quesada, M., Korb, F. M., Funes, M. J., Lupiáñez, J., & Egner, T. (2014). Comparing neural substrates of emotional vs. non-emotional conflict modulation by global control context. *Frontiers in Human Neuroscience*, *8,* 66. http://doi.org/10.3389/fnhum.2014.00066

Tortosa, M. I., Strizhko, T., Capizzi, M., & Ruz, M. (2013). Interpersonal effects of emotion in a multi-round Trust Game.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., … Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. http://doi.org/10.1016/j.psychres.2008.05.006

Vogt, B. A., Finch, D. M., & Olson, C. R. (1992). Functional heterogeneity in cingulate cortex: the anterior executive and posterior evaluative regions. *Cerebral Cortex (New York, N.Y.: 1991)*, *2*(6), 435–443.

Watanabe, T., Yahata, N., Kawakubo, Y., Inoue, H., Takano, Y., Iwashiro, N., … Yamasue, H. (2014). Network structure underlying resolution of conflicting non-verbal and verbal social information. *Social Cognitive and Affective Neuroscience*, *9*(6), 767–775. http://doi.org/10.1093/scan/nst046

Watson, R., Latinus, M., Noguchi, T., Garrod, O., Crabbe, F., & Belin, P. (2013). Dissociating task difficulty from incongruence in face-voice emotion integration. *Frontiers in Human Neuroscience*, *7*, 744. http://doi.org/10.3389/fnhum.2013.00744

Winston, J. S., Henson, R. N. A., Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-Adaptation Reveals Dissociable Neural Representations of Identity and Expression in Face Perception. *Journal of Neurophysiology*, *92*(3), 1830–1839. http://doi.org/10.1152/jn.00155.2004

Wittfoth, M., Küstermann, E., Fahle, M., & Herrmann, M. (2008). The influence of response conflict on error processing: evidence from event-related fMRI. *Brain Research*, *1194*, 118–129. http://doi.org/10.1016/j.brainres.2007.11.067

Zaki, J., Hennigan, K., Weber, J., & Ochsner, K. N. (2010). Social Cognitive Conflict Resolution: Contributions of Domain-General and Domain-Specific Neural Systems. *The Journal of Neuroscience*, *30*(25), 8481–8488. http://doi.org/10.1523/JNEUROSCI.0382-10.2010

Capítulo 6: Serie Experimental IV. Emotional conflict: interactions between personal identity and emotional expression in predictive judgments as shown by ERPs

El contenido de este capítulo está en preparación como Alguacil, S., Madrid, E., Tudela, P., & Ruz, M. Emotional conflict: interactions between personal identity and emotional expression in predictive judgments as shown by ERPs.

Abstract

Two of the most relevant sources of information to guide decisions are the behavioral tendencies associated to specific personal identities and their facial emotional displays. The present electrophysiological experiment aimed to study the stages of face processing enhanced by the use of these cues and their interactions. Participants played a modified Trust Game with 8 different alleged partners, and in separate blocks either the identity or the emotions carried information regarding the partner's proximate behavior. Behavioral responses showed that whereas decisions were equal across blocks, participants were faster when they used the identity compared to the emotional expression of the partners. We also observed that in blocks where emotions were irrelevant and non-predictive, they still influenced the speed of choices depending on the congruency between the expectations associated to the identity and the emotion displayed. Electrophysiological results showed early task effects (P1) and differences between faces associated to gain and loss, but only for predictions based on emotions (N1 and VPP). On the other hand, only predictions based on personal identity modulated the N200, a potential that has been linked to the analysis of value of action for guide decisions, and diminished the amplitude of the P3b. In addition, the interference effect generated by the emotions ignored during the identity block was reflected on the N170 amplitude, as expressions of anger enhanced

the amplitude of this potential but only for partners predicting gains, and the P3b potentials. Overall, our results suggest that using the identity or the emotion displayed by a face to predict cooperation tendencies recruits dissociable neural circuits from an early point in time, and that both sources of information generate interactive early and late patterns.

6.1. Introduction

Faces represent complex visual patterns that convey a wide array of information about individuals such as their identity, facial emotional expression, sex and age (Bruce & Young, 1986; Palermo & Rhodes, 2007; Pizzagalli et al., 2002). Human perceivers display a high degree of competence at extracting this information in a fast and efficient manner. Thus, faces represent a ‘special’ kind of stimuli, so that they are perceived and processed differently than other objects since early in life (de Haan, Pascalis & Johnson, 2002; Nelson, 2001). A few hundred milliseconds suffice to extract both changeable and invariant facial features that can be used to guide our future behavior toward others (Bruce & Young, 1986; Ekman & Friesen, 2015; Fridlund, 2014; Haxby, Hoffman, & Gobbini, 2000; Oosterhof & Todorov, 2008; Willis & Todorov, 2006). Among these, facial identity and emotional expression are essential to social adaptation (Frijda, 1988; Frijda & Mesquita, 1994; Todorov, Said, Engell, & Oosterhof, 2008).

The ability to extract information from faces is supported by a specialized and distributed decoding neural network composed by a Core System of visual extrastriate areas that acts in concert with complementary regions of an Extended System, to achieve a full decoding of different aspects of faces (Haxby & Gobbini, 2011; Haxby et al., 2000). The Core

System, involved in processing the visual appearance of faces, is composed by different regions which, given their hierarchical convergent structure, provide input to subsequent areas (Haxby & Gobbini, 2011; Haxby et al., 2000). This core involves the inferior occipital gyrus, which processes face-part information at an early stage (Haxby et al., 2000; Pitcher, Walsh, & Duchaine, 2011). Such information is provided as input to both the lateral fusiform gyrus and superior temporal sulcus (Haxby et al., 2000). While the fusiform gyrus (FG; also referred as face fusiform area, FFA in the face processing literature; Kanwisher, McDermott, & Chun, 1997) is involved in the processing of invariant facial features (i.e. identity), the superior temporal sulcus (STS) decodes changeable facial aspects as the emotional expressions (Haxby & Gobbini, 2011; Haxby et al., 2000). Finally, other limbic areas, such as the amygdala, have been included in the Core System given their role in emotional decoding (Haxby & Gobbini, 2011; Haxby et al., 2000). In coordination with core areas, regions from the Extended System represent the meaning of facial configuration (Haxby et al., 2000). Several regions, including the anterior paracingulate cortex, the temporoparietal junction, the anterior temporal cortex, the precuneus and the posterior cingulate, participate in the attribution of meaning of personal knowledge linked to invariant features (Gobbini & Haxby, 2007). In addition, the insula and striatum are involved in decoding of the meaning of emotional content linked to changeable features (Gobbini & Haxby, 2007). Other classical models in face perception (Bruce & Young, 1986) also propose that the extraction of identity and emotional information from faces takes place through different means.

There are, however, discrepancies as to whether each type of features is processed by fully separated regions or they interact to some extent (Baudouin, Gilibert, Sansone, & Tiberghien, 2000; Bruce & Young, 1986; Haxby & Gobbini, 2011; Haxby et al., 2000). Haxby and colleagues (2000), propose that perceptual operations on invariant and changeable facial aspects, given their potential to interfere with each other, are kept distinct within the system by anatomic segregation of their representations (Haxby & Gobbini, 2011). This anatomic and functional dissimilitude is partially justified by the fact that changes in emotional expression do not influence the decoding of identity, whereas the same emotional expression has a similar meaning across different individuals (Haxby & Gobbini, 2011). Some studies, on the other hand, have provided data suggesting that both facial features do influence each other in certain circumstances. For example, Schweinberger & Soukup (1998) investigated whether variations in an irrelevant stimulus dimension on judgments of faces influenced the responses to a relevant identity, emotional expression, and facial speech dimensions. They showed that identity judgments were unaffected by the other two dimensions, but during emotion or facial speech judgments both dimensions were influenced by identity. Hence, identity may influence the processing of emotion whereas emotion *per se* cannot change the way in which identity is processed (Schweinberger & Soukup, 1998). Similarly, a more recent study showed that emotional expression influenced the degree at which faces were categorized as familiar or unfamiliar (Baudouin et al., 2000). More interestingly, smiling faces increased the degree of familiarity perceived for both famous and unknown identities (Baudouin et al., 2000). This identity-emotion asymmetry is also questioned by evaluations of trustworthiness, as claimed by the over-generalization hypothesis (Todorov, Mandisodza, Goren, & Hall, 2005; Todorov et al., 2008). The evaluation of trustworthiness of emotionally neutral faces, which implies identity recognition, seems to be an over-generalization of functionally adaptive mechanisms for rapidly detecting emotions in other people (Todorov et al., 2005, 2008). This over-generalization appears to be grounded on basic facial features that resemble emotional expressions signaling approach/avoidance tendencies (Todorov et al., 2008; Todorov & Uleman, 2002). Thus personal identity, at least in part, is constructed using the emotional content of faces. This would support the hypothetical existence of some sort of mutual dependency between invariant and changeable facial features.

How identity and emotional expression are processed under the influence of each other is of paramount importance not just because an accurate recognition is needed, but also since both facial features serve to signal the probable behavior underlying intentionality during personal interactions (Frijda, 1988; Frijda & Mesquita, 1994; Oosterhof & Todorov, 2008; Todorov et al., 2008; Tooby & Cosmides, 1990). Crucially, intentionality inferred from invariant and changeable facial aspects may be contradictory. For example, attributions of intentionality based on identity (e.g. someone is described as trustworthy) may create expectations of positive future outcomes (Oosterhof & Todorov, 2008; Todorov et al., 2008), but if the same individual looks angry, then intentions naturally associated with emotional facial content generate an expectation of negative outcomes as a result of his/her behavior (Frijda, 1988; Frijda & Mesquita, 1994). Consequences would be identical for the inverse situation (i.e. someone described as untrustworthy but displaying happiness). From this perspective, a group of studies has explored the strategic use of facial identity and emotional expressions during social decision-making scenarios (Alguacil, Tudela, & Ruz, 2015; Tortosa, Lupiáñez, & Ruz, 2013; Tortosa, Strizhko, Capizzi, & Ruz, 2013). When emotions are explicitly employed to predict other people's behavior, participants need less time to familiarize with their natural associations (i.e. 'happiness-trustworthy' and 'anger-untrustworthy') than if they are asked to assign unnatural expectations to the emotional content of faces (Tortosa, Strizhko, et al., 2013). In addition, when participants are required to attend to facial identity, defined as the relevant cue, while the emotional expression is explicitly described as non-predictive and asked to be ignored, the automatic interference of the emotional content seems to be unavoidable despite its irrelevance (Alguacil et al., 2015; Tortosa, Lupiáñez, et al., 2013; Tortosa, Strizhko, et al., 2013).

The evidence mentioned so far suggests that emotional information generates an unavoidable influence during identity judgments. Given the rapid processing of identity and emotional information (Barrett, 2012; Barrett & Bliss-Moreau, 2009), the emotional influence could take place from the early visual inspection of facial elements. The description of the temporal dynamics of these processes is essential to understand how humans extract identity and emotional expression from faces (Batty & Taylor, 2003; Bentin, Allison, Puce, Perez, & McCarthy, 1996; Pizzagalli et al., 2002). Under this rationale, the aim of the present study was to employ electroencephalographic measurements to evaluate how the strategic use of identity and emotional expression as relevant cues for guiding decisions influences different stages of information processing. In addition, we studied the levels at which irrelevant emotions influence the evaluation of facial identity.

Previous literature highlights several evoked potentials (ERP) that could be modulated by the use of identity or emotion as cues to predict outcomes and also by conflict between them. The earliest ERP responses to face stimuli are usually observed approximately one hundred milliseconds after face onset at the posterior bilateral P100 and at the fronto-central N1 (Dennis, Malone, & Chen, 2009; Eimer & Holmes, 2002; Hilimire, Mienaltowski, Blanchard-Fields, & Corballis, 2014; Wong, Fung, McAlonan, & Chua, 2009; Yang, Gu, Guo, & Qiu, 2011). At this level of processing, previous research has reported early orientation of attention to different social categories extracted from faces (i.e. black versus white people; e.g. Ito & Urland, 2003). Similarly, it has been suggested that these potentials may reflect a fast global extraction of emotional salient information before a more detailed face decoding of fine-grained information takes place (Dennis et al., 2009; Eimer & Holmes, 2002; Hilimire et al., 2014; Pitcher, Walsh, Yovel, & Duchaine, 2007; Pizzagalli, Lehmann, et al., 2002; Pourtois, Dan, Grandjean, Sander, & Vuilleumier, 2005; Pourtois, Grandjean, Sander, & Vuilleumier, 2004; Vuilleumier & Pourtois, 2007). Hence, we expect amplitude variations at this processing stage as a result of attention to emotional information, but not by facial identities, as stimuli were equated at their low-level characteristics. Emotional capture could be possibly reflected in a heightened amplitude for negative expressions as they seem to naturally grab attention as shown, for example, by the early ‘negative bias’ (Carretié, Hinojosa, Martín-Loeches, Mercado, & Tapia, 2004; Eimer, 2011; Eimer & Holmes, 2002; Vuilleumier & Pourtois, 2007). Of special interest is the N170, with posterior bilateral distribution, and the Vertex Positive potential (VPP), with a central location, which appears around one hundred and seventy milliseconds after face onset (Bentin et al., 1996; Rossion et al., 2000; Rossion, Joyce, Cottrell, & Tarr, 2003), although earlier timings have been reported (Rossion & Caharel, 2011). Both potentials, which seem to underlie similar mental operations (Eimer, 2011; Joyce & Rossion, 2005; Yang et al., 2011), have been linked to the decoding of more complex configural facial aspects (Eimer, 2011; Jeffreys, 1989; Joyce & Rossion, 2005; Rossion et al., 2000). They are sensitive to the categorization of facial identity and also of emotional expressions (Eimer, 2011; Vuilleumier & Pourtois, 2007). Even more, previous evidence has shown interacting effects between the processing of emotion and race as reflected on the amplitude of the N170; emotional expressions seem to differently affect its amplitude depending on the race of the individual that expresses them (Tortosa, Lupiáñez, et al., 2013). A similar interacting effect has been found in an emotional face–word Stroop task from a previous study where participants were required to respond to emotional expression or word meaning (Zhu, Zhang, Wu, Luo, & Luo, 2010). Authors found an enhanced N170 during the emotional expression task, when face and word had incongruent valence. Others previous studies have found congruency effects at this stage between different elements containing affective information (see for a review about this issue Hinojosa, Mercado, & Carretié, 2015). Therefore, we expected to find an interacting effect of these two facial features, with a more pronounced amplitude for incongruent emotion/identity valence value combinations (Zhu et al., 2010).

The N2 potential appears from 200 to 300 ms after stimulus onset at central or fronto-central locations, and has been related to cognitive control and conflict processing (e.g. Folstein & Van Petten, 2008; Heil, Osman, Wiegelmann, Rolke, & Hennighausen, 2000; Kopp, Rist, & Mattler, 1996). This potential seems sensitive to the detection of incongruence between alternatives (Folstein & Van Petten, 2008). Although it has been extensively probed its sensitivity to conflict when non-emotional elements lead to incongruent representations (Folstein & Van Petten, 2008) and to the influence of affective background during the resolution of cognitive conflict (Kanske & Kotz, 2010a, 2010b, 2011), other paradigms employing socially interactive contexts do not show emotional conflict at this stage of information processing (Alguacil, Tudela, & Ruz, 2013; Ruz et al., 2013; Tortosa, Lupiáñez, et al., 2013; Panadero & Tudela, *in preparation*). Previous studies have shown that the N2 is also sensitive to the processing of social expectations extracted from identity (Derks, Stedehouder, & Ito, 2015; Tortosa, Lupiáñez, et al., 2013) and from expressed emotions (Derks et al., 2015; Ruz et al., 2013; Tortosa, Lupiáñez, et al., 2013). In a similar temporo-spatial location to that of the N2 potential, by employing a modified version of the Trust Game (TG), Ruz and colleagues (2013), found an interaction between executive attention an emotion for trustworthy identities, which revealed a heightened amplitude for trustworthy identities with an expression of anger (Ruz et al., 2013). Based on these evidences, we hypothesized that the amplitude of the N2 potential would be heightened for stimuli with a negative value respective to the predicted outcome.

At later processing stages, the P3b appears in a 300-600 ms post-stimulus range at centro-parietal electrodes (Polich, 2007). This potential is influenced by facial identities and emotions (see for example Campanella et al., 2000; Campanella, Quinet, Bruyer, Crommelinck, & Guerit, 2002). Specifically, variations in its amplitude have been associated with the discrimination of different emotional expressions (Luo, Feng, He, Wang, & Luo, 2010; Tortosa, Lupiáñez, et al., 2013) and facial predictive identities of trustworthy or untrustworthy behaviors (Ruz et al., 2013). Its amplitude also reflects changes as a result of outcome evaluation and reward processing in social interactions during economical exchanges ,as for example in gambling task (Wu & Zhou, 2009). In this line, stimuli with high motivational significance, as for example those that signal personal benefit and gains, seem to enhance the amplitude of this potential (Moser, Gaertig, & Ruz, 2014; Ruz et al., 2013; Yeung & Sanfey, 2004). The P3b has been largely related to the inhibition of 'irrelevant' information and also to the deployment of attentional resources on the 'relevant' one (Polich, 2007). The harder the implementation of inhibition mechanisms, the smaller the amplitude of the P3b (Polich, 2007). Thus, the P3b amplitude is usually larger for congruent than for incongruent trials (Neuhaus et al., 2010; Valle-Inclán, 1996). Hence, we expected an enhancement of the P3b for motivationally salient information signaling future positive outcomes, but also, in line with previous conflict and outcomes/expectancies evaluation literature (Neuhaus et al., 2010; Valle-Inclán, 1996), we hypothesized that the P3b amplitude would be reduced in incongruent situations where expectancies from identity and emotion do not match.

In summary, although previous studies have explored how identity and emotion are decoded, these have mostly employed simple categorization tasks (Atkinson, Tipples, Burt, & Young, 2005; Schweinberger & Soukup, 1998), or paradigms where only emotion or identity were the relevant elements for predicting outcomes (Alguacil et al., 2015; Ruz et al., 2013; Ruz & Tudela, 2011; Tortosa, Lupiáñez, et al., 2013). Even more, most studies to date have employed paradigms fully devoid of social context, which represents a drawback given the innate social nature of emotional phenomena (e.g. Parkinson, 1996). Bearing all this in mind, in the present study, we employed three modified versions of the Trust Game, frequently employed to explore mental operations during social interactions in a more natural setting, where participants earned money by predicting the most likely behavior of partners, represented by facial photographs displaying emotional expressions. In separate blocks, they had to use the identity, the emotion or the color of a frame bordering the target pictures (control condition) as cues to predict the most likely trial outcome. At the beginning of each block, participants were informed of which of the three cues was predictive (83% of validity) of whether the trial would return point earnings or not (the other two irrelevant cues remained unpredictive). Each cue was predictive in one of the three game versions, and explicitly defined as relevant for anticipating outcomes; in the other two blocks, the cue was defined as non-predictive and non-relevant for the task. Participants were paid in order to increase their motivation to follow predictions guided by facial information about outcomes at the end of the trial and also with the purpose to create a more realistic monetary exchange. Considering the evidence mentioned evidence, at the behavioral level, although we expect the relevant and predictive facial features to guide decisions so participants would invest a higher number of times their money when expected for gain than for loss, we also hypothesized that participants would need more time to make their decisions in situations where the ignored emotional features lead to incongruent predictions about outcomes in the trial. This procedure, in combination with electrophysiological measures, will allow for an equivalent evaluation of identity and expression decoding so we will be able to contrast the employment of both features when they act as relevant or irrelevant cue for future outcomes in a more natural, than traditional stimulus-response paradigms and when high order mental abilities are required. Such manipulation will also permit to explore their potential interaction and tracing the processing level at which it can occur.

6.2. Methods

6.2.1. Participants

Twenty-eight healthy volunteers were recruited from the University of Granada in exchange for course credits. All of them had normal or corrected-to-normal vision. Twelve were female, their mean age was 23 years (range: 18-40) and two of them were left-handed.

In addition to course credits, participants received payment according to their earnings during the experimental task (ranging between 3.5 and 7.5 EUR). All participants signed a consent form approved by the local Ethics Committee.

6.2.2. Task

Participants played a game composed by three different multiple-round adaptations of the Trust Game (Berg, Dickhaut, & McCabe, 1995; Ruz & Tudela, 2011) presented in three separate blocks: Emotion, Identity and Color. In all of them, participants received 1 EUR at the beginning of each trial. Then they decided whether to keep it (which resulted in a gain of 1 EUR) or to bet it. Different cues in each block predicted the most likely trial outcome. In the Emotion block, the photographs of four happy or angry faces cued a most likely gain/loss, respectively. In the Identity block, two facial identities cued a likely gain and the other two a likely loss. In the Color block, the color of a frame surrounding pixelated faces served as the cue for gain or loss (see Fig. 1). These relevant cues were valid on 83% of the trials. The non-predictive features (identity and color in the Emotion block, emotion and color in the Identity block, and identity and emotion in the Color block) were non-predictive (50%).

Participants were instructed to respond during the time the facial display was on the screen by pressing one of two buttons with their left or right index fingers. In all cases, if participants decided to bet, the initial sum was multiplied by five (5 EUR) and a feedback symbol informed them of the gain (2.5 EUR) or loss (0 EUR).

6.2.3. Stimuli and procedure

Forty-eight faces (12 identities, 6 females) displaying happy or angry (half of them with the mouth open) emotional expressions were taken from the NimStim set (Tottenham et al., 2009). Two different female and two male identities were used for each of the blocks (counterbalanced across participants). Adobe Photoshop CS 6 was employed to pixelate the original pictures, (5×5 pixels; Adobe Systems Inc., San Jose, California, USA) for the color version of the game. Pictures were framed by colored squares in all blocks. In total, six colors were employed for this purpose, in 2 different tones each (dark and light blue, green, orange, purple, and yellow). Pairs of two different colors were used in each block (counterbalanced across participants).

The order of the blocks, the association between hand and response, the feedback symbols and their color were fully counterbalanced across participants. The task was implemented using E-Prime 2.0 Professional software (Schneider, Eschman, & Zuccolotto, 2002). All stimuli were centrally displayed in a 17-inch CRT monitor, against a grey background (see Fig. 1). A trial started with a fixation cross lasting 3350 ms on average (random 3100-3600 ms; +; 0.5º on average), followed by a framed picture of a face (7.15° on average) for 1500 ms, another fixation cross for 200 ms on average (random 100-300 ms), and finally the feedback symbol ('*', 0.67°, or '#', 0.57° on average) for 200 ms. On average, a trial lasted for 5250 ms. Each participant completed a total of 576 experimental trials (192 for each of the three blocks) for a total of about 55 minutes. At the beginning of the session, participants responded to 60 practice trials (20 per block); during this practice the cue was always followed by a valid feedback symbol.

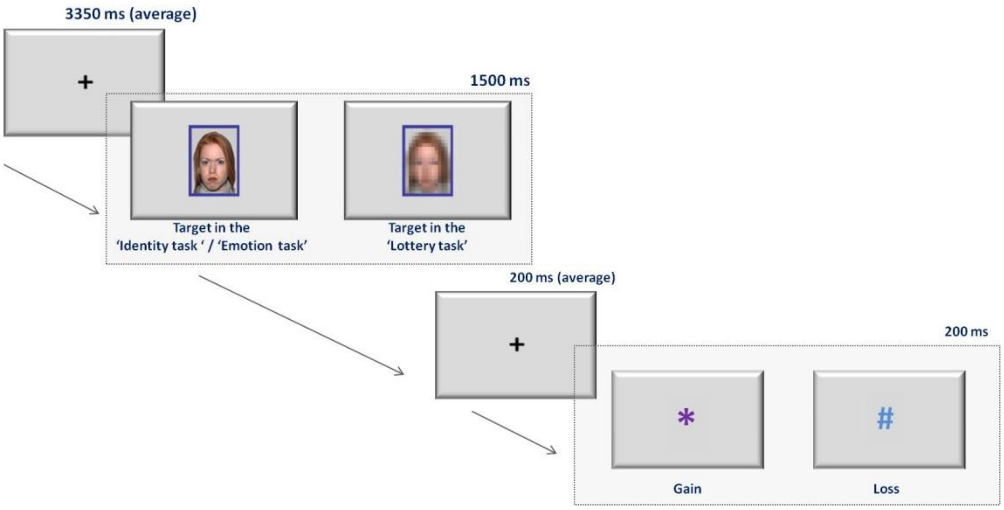

Figure 1. Sequence of events in a trial.

6.3. Electrophysiological recording and analysis

Participants were seated in front of the computer monitor in an electrically shielded room and were instructed to avoid eye blinks and movements during stimulus presentation and responses. EEG was recorded with a high-density 128-channel EEG system Geodesic Sensor Net (Tucker, 1994), referenced to the vertex channel. The head coverage included sensors lateral to and below both eyes to monitor horizontal and vertical eye movements (HEOG and VEOG). The EEG net was connected to an AC-coupled, high-input impedance amplifier (200 MΩ). At the beginning of the recording session, impedance was measured for each channel and was set below 50 kΩ, as recommended for Electrical Geodesics high-input impedance amplifiers. Amplified analog voltages (0.1—200-Hz band pass) were digitized at 1000 Hz[4] (12 bits A/D converter and 0.02 μV minimum resolvable voltage). The continuous EEG was filtered offline using a 40 Hz low-pass filter. After that, the EEG was segmented 200 ms before and 800 ms after target onset and processed for artifact detection. Trials containing eye blinks or eye movements (electro-oculogram channel differences greater than 70 μV) or more than 20% of bad channels were excluded. Data from bad channels were later replaced using a spherical interpolation algorithm (Perrin, Pernier, Bertrand, & Echallier, 1989). ERPs were re-referenced offline to the average. A 200 ms pre-stimulus interval was used as baseline.

A minimum criterion of 30 artifact-free trials per subject and condition was established to maintain an acceptable signal-to-noise ratio. Voltage analyses were performed on the spatio-temporal windows that captured the grand-

[4] Data were later subsampled from 1000Hz to 250Hz.

average peaks of the P100, N1, N170, VPP, N200 and P3b potentials. The selected electrodes were those were the components were maximally distributed (see Figs. 2, 4, 5 and 6). The temporal windows chosen were centered on the peak of the potentials in the grand-averaged waveforms.

Trials without a response were not considered either in the behavioral or in the ERP analysis. Four participants were excluded from both because of excessive artifacts during recording. An initial observation of the grand-average data revealed that the ERPs of the Color block showed marked differences in latency and shape with the other two blocks, most likely mainly due to the pixelation of the faces (see Fig. 1). Because of this, and given that our main hypothesis concerned the comparison of the Emotion and Identity blocks, data from the Color block were only considered in the behavioral analysis.

To address our research questions, two separate analyses were carried out. The first one, which will be referred as 'overall analysis', analyzed the key stages that varied according to whether facial displays of emotion or personal identities were used as cues to predict a future economic outcome. To do this, the mean amplitude of face-locked ERPs, averaged over the selected channels and time windows, were submitted to repeated-measures ANOVA including the Block (Emotion vs. Identity) and Prediction (Gain vs.

Loss) as factors. To make data across blocks comparable, of the 4 conditions that participants received in the Identity Block (Gain/Happy, Gain/Angry, Loss/Happy, and Loss/Angry), the two that were not presented in the Emotion block to avoid predictive conflict (Gain/Angry and Loss/Happy, see Ruz & Tudela, 2011) were not included in the average. The number of observations was equated across conditions (45 trials on average).

The second analysis, from now on denominated as 'emotional conflict analysis', sought to evaluate the stages of processing affected by the unavoidable emotional conflict. This set of ANOVAs employed the same electrode montages and time windows as before but only with data from the Identity block (in which emotions were ignored). Factors entered were Prediction (Gain vs. Loss) and the Facial Expression (Happy vs. Angry). To avoid unnecessary duplication of results, for this analyses we only focused on the interaction between the Prediction and Facial Expression factors, so no other effects will be reported.

Given the relevance of lateralization for face processing, as reported by previous evidence (Campanella et al., 2000; Luo et al., 2010; Pizzagalli et al., 2002; Wager, Phan, Liberzon, & Taylor, 2003), the Hemisphere was included as relevant factor for the P100, the N170 and the P3b potentials in both analyses.

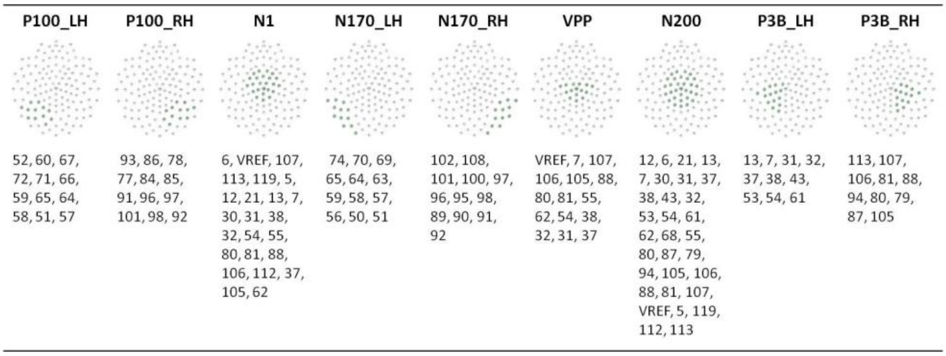

Figure 2. Spatial location and numbers of electrodes used for the ERP analyses. LH: Left Hemisphere; RH: Right Hemisphere.

6.4. Results

6.4.1. Behavioral

On average, participants betted their money on 49% of the trials. As revealed by the overall analysis, participants mean bet rates yielded a main effect of Prediction, F (1, 23) = 442.41, p < .001, $\eta p2$ =.95, as they risked more money when they anticipated gain (M = 0.88, SE = 0.28) than when they expected loss (M = 0.10, SE = 0.16). For the Emotional Conflict analysis, besides the Prediction effect, the effect of emotion approached significance levels, F (1, 23) = 3.921, p = .06, $\eta p2$ =.146, with a tendency for participants to bet money slightly more frequently with happy, (M = 0.51, SE = 0.02), than with angry partners, (M = 0.48, SE = 0.01).

The overall analysis[5] on the participants' mean reaction times (RTs) revealed a significant main effect of the Block and Prediction factors. For the Block, F (1, 23) = 22.200, p < .001, $\eta p2$ =.67, participants were slower responding in the Emotion (M = 739.87 ms, SE=20.05), than in the Identity block (M = 701.11 ms, SE = 17.63) or in the Color one (M = 638.17 ms, SE = 15.81), F (1, 23) = 9.665, p < .01, $\eta p2$ =.30 and F (1, 23) = 45.548, p < .001, $\eta p2$ =.66, respectively. The difference between the Identity and the Color blocks also reached significance, F (1, 23) = 23.553, p < .001, $\eta p2$ =.51. Responses were slower for expectations of loss, M = 705.33, SE = 15.73, than gain, M = 680.76, SE = 17.28. F (1, 23) = 11.21, p < .01, $\eta p2$ =.33. No other effect or interactions were significant, all Fs < 1. The ANOVA for the Emotional Conflict on the mean RTs of the Identity block showed the expected interaction between Prediction and Facial Expression, F (1, 23) = 13.42, p = .001, $\eta p2$ =.37. Planned comparisons showed that participants were slower to respond to a partner whose identity cued probable gain but that displayed anger, (M = 699.34, SE = 19.52), than when the same identity showed an expression of happiness, (M = 675.05, SE = 17.86). No other effect or interactions were significant; all ps > .290 (see Fig. 3).

[5] A separated analysis including the Block (Emotion, Identity or Color) and the Prediction (Gain, Loss) for the same factors levels than in the ERP analysis did not change the general pattern of results.

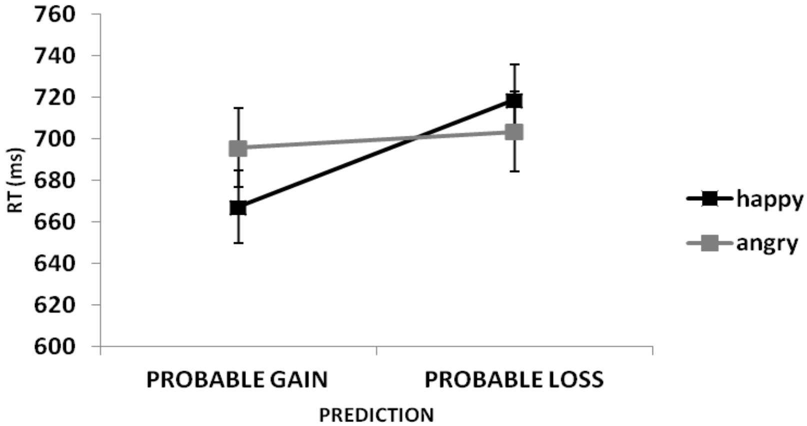

Figure 3. Emotional conflict when participants predicted gain or loss based on facial identity. Error bars represent standard error of the mean.

6.4.2. Electrophysiological

P100

The P100 peaked at around 100 ms and was analyzed from 80 to 120 ms over bilateral posterior electrodes (see Figs. 2 and 4). The ANOVA yielded a main effect of Hemisphere, F (1, 23) = 5.018, $p < .05$, $\eta p2$ =.18, as this potential was more positive at right (2.17 μV; SE = 0.36) than at left hemisphere electrodes (1.50 μV; SE = 0.30). The interaction between Block and Hemisphere was close to significance, F (1, 23) = 4.212, p = .052, $\eta p2$ =.15. The Hemisphere was significant for the Emotion block, F (1, 23) = 7.553, $p < .05$, $\eta p2$ =.25, with larger positive voltages over right (2.28 μV; SE = 0.36) than over left electrodes (1.43 μV; SE = 0.31). This effect was not significant for the Identity block, F = 2.365, p = .138.

The interaction between Prediction and Emotion in the Emotional Conflict ANOVA was not significant ($F < 1$).

N1

The N1 over fronto-central electrodes peaked at 100 ms and was analyzed from 80 to 120 ms (see Figs. 2 and 4). The overall ANOVA showed a main effect of Prediction, better explained by its interaction with the Block factor, F (1, 23) = 9.083, $p < .01$, $\eta p2$ =.28. Planned comparisons contrasts revealed an effect of Prediction close to significance for the Emotion block, F (1, 23) = 3.794, $p = .06$, $\eta p2$ =.14. The N1 was more negative for happy faces, which cued gain (-1.11 μV; SE = 0.20) than for angry ones, associated with probable loss (-0.82 μV; SE = 0.20). The same contrasts were not significant in the Identity block, F = 2.447, p = .131.

The interaction between Prediction and Emotion in the Emotional Conflict ANOVA was not significant ($F < 1$).

N170

This potential peaked approximately at 150 ms after face onset and was analyzed over posterior bilateral electrodes (see Figs 2 and 4a) from 125 to 175 ms. The overall ANOVA revealed a significant interaction between Block and Hemisphere, F (1, 23) = 5.931, $p < .05$, $\eta p2$ =.20. Differences in amplitude between the Emotion and the Identity blocks were significant for the left, (Emotion: -2.59 *μV*; *SE* = 0.85; Identity: -2.28 *μV*; *SE* = .40), F (1, 23) = 6.031, $p < .05$, $\eta p2$ =.21, but not for the right hemisphere, $F < 1$.

The Emotional Conflict analysis revealed a significant interaction between Facial Expression and Prediction factors, F (1, 23) = 5.345, $p < .05$, $\eta p2$ =.19. Subsequent analysis showed that when identities predicted gain, the N170 was more negative for expressions of anger than for those of happiness (Angry: -2.84 *μV*; *SE* = 0.40; Happy: -2.19 *μV*; *SE* = 0.42), F (1, 23) = 11.804, $p < .01$, $\eta p2$ =.34. There were no significant differences for identities that predicted loss, F = 1.436, p = .243 (see Fig. 4b).

VPP

With identical temporal distribution than the N170 but over central electrodes (see Figs. 2 and 5), the overall ANOVA showed an interaction between Block and Prediction, F (1, 23) = 7.217, $p < .05$, $\eta p2$ =.24. Whereas in the emotion block the VPP was more positive for angry expressions (1.58 *μV*; *SE* = 0.30) associated with loss, than for happy ones (2.06 *μV*; *SE* = 0.34) that cued gain, F (1, 23) = 7.902, $p < .05$, $\eta p2$ =.26, this was not the case in the Identity block, F = 1.034, p = .320.

The interaction between Prediction and Emotion in the Emotional Conflict analysis was not significant ($F < 1$).

N200

The N200 peaked at 230 ms and was analyzed from 200 to 260 ms over central electrodes (see Figs. 2 and 5). The overall analysis revealed a significant interaction between Block and Prediction, F (1, 23) = 9.820, $p < .01$, $\eta p2$ =.30. Posterior analyses showed a main effect of Prediction for the Identity block, F (1, 23) = 5.102, $p < .05$, $\eta p2$ =.18, due to a larger negative amplitude for faces associated to future loss (-0.28 μV; SE = 0.25) than for those that cued gain (0.21 μV; SE = 0.26). This was not the case for the Emotion block, F = 2.320, p = .141.

The interaction between Prediction and Emotion in the Emotional Conflict ANOVA was not significant ($F < 1$).

P3b

The P3b potential peaked at 475 ms and was analyzed from 445 to 515 ms over bilateral central electrodes (see Figs. 2 and 6a). The overall ANOVA showed a main effect of Hemisphere, F (1, 23) = 4.457, $p < .05$, $\eta p2$ =.16, as this potential was more positive at right (2.63 μV; SE = 0.26) than at left electrodes (2.21 μV; SE = 0.27). There was also a main effect of Prediction, F (1, 23) = 4.392, $p < .05$, $\eta p2$ =.16, better explained by its interaction with the Block factor, F (1, 23) = 15.014, p = .001, $\eta p2$ =.39 (see Fig. 6a). Subsequent comparisons revealed a main effect of Prediction for the

Identity block, F (1, 23) = 13.802, p = .001, *ηp2* =.37, as the P3b was more positive for faces linked to gain (2.74 *μV*; *SE* = 0.31) than to loss (2.04 *μV*; *SE* = 0.27**).** This effect was not significant for the Emotion block, $F < 1$.

The Emotional Conflict analysis yielded an interaction between Hemisphere, Prediction and Facial Expression, F (1, 23) = 6.538, $p < .05$, *ηp2* =.22. There was a main effect of Prediction for right hemisphere electrodes, F (1, 23) = 4.797, $p < .05$, *ηp2* =.17, identical to that founded in the overall ANOVA. At left locations, Prediction and Facial Expression interacted, F (1, 23) = 8.172, $p < .01$, *ηp2* =.26. For identities predicting gain, the P3b was larger for happy (2.22 *μV*; *SE* = 0.37) than for angry (1.60 *μV*; *SE* = 0.32) facial expressions, F (1, 23) = 9.481, $p < .01$, *ηp2* =.29. The emotional expressions of identities predicting loss, on the other hand, did not modulate the P3b, $F<1$ (see Figs. 6b and 6c).

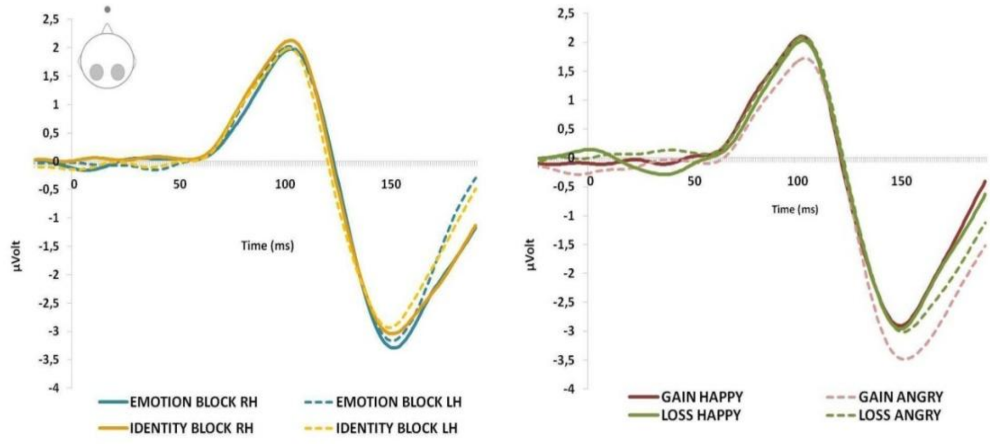

Figure 4: Face-locked ERPs showing the modulation of the N170 potential by the effect of the Block at the left hemisphere (4a: left panel) and by the Facial Expression on the Identity block for faces that mostly predicted Gain (4b: right panel).

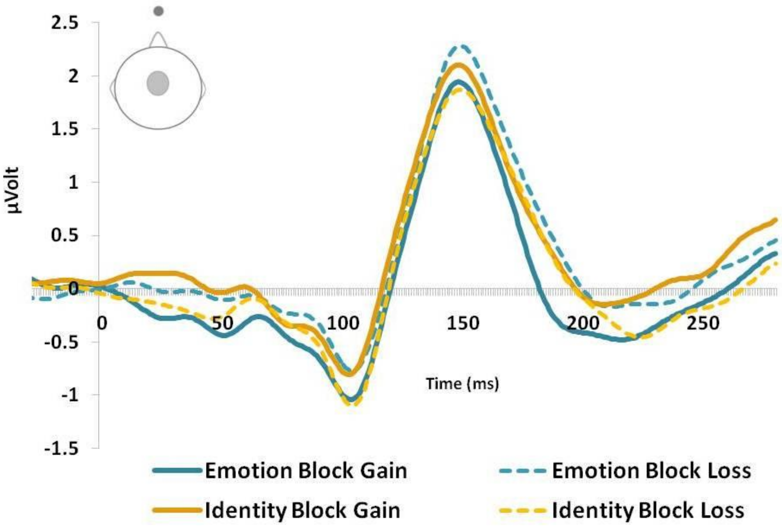

Figure 5. Face-locked ERPs showing the modulation of the fronto-central N1 and central VPP by the Emotion as relevant cue and by the central N200 by the Identity as predictive cue. The spatial window employed for the analyses

is represented in the upper-left diagram. Positivity is plotted upwards in all figures.

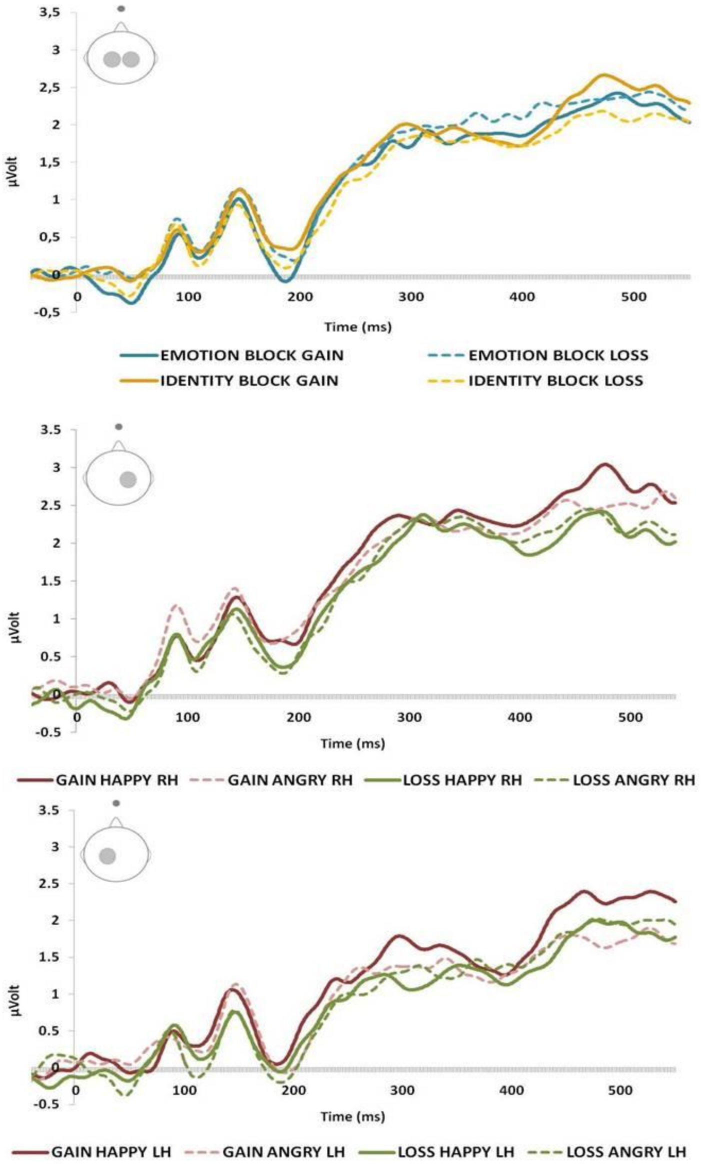

Figure 6: Face-locked ERPs showing the effect of Prediction on the Identity task in the overall analysis (6a: top pannel) and also lateralized to the right hemisphere as shown by the Emotional Conflict ANOVA (6b: middle pannel) on the P3b. The display also shows the conflict effect of Emotion on the Identity task at the left hemisphere on the P3b (6c: bottom pannel).

6.5. Discussion

The aim of the present study was to evaluate how the strategic use of facial identity and emotional expression influences different processing levels when they are the relevant cues for the anticipation of future economic consequences during a decision-making scenario. We also evaluated how and at which processing levels ignored and non-predictive facial emotions interact with identity decoding. Overall, our behavioral results showed that participants used the relevant information to predict near outcomes and make their decisions, which took longer for emotion than for identity judgments. In addition, decisions were slower when the ignored emotions led to expectations incongruent to the outcomes predicted by personal identity information. Thus, contradictory predictions generated by the facial expression generate an unavoidable conflict during facial identity processing. Our electrophysiological data suggest early hemispheric differences according to the cue employed to draw predictions, and also rapid perceptual interactive adjustments depending on the nature of this cue and the valence of the prediction. Whereas predictions driven by emotions enhanced the fronto-central N1 and VPP potentials, identity predictions modulated the later N2 and P3b. In addition, the conflict stemming from ignored emotions during the identity-based predictions influenced initial levels of face decoding, as reflected on the N170, and decision-related processes indexed by the P3b.

The behavioral strategy followed by participants during the game, to decide whether or not to bet their money, is coherent with the role of motivational factors in the consecution of rewards. Previous studies have shown that expectations about the amount of reward, or its probability, influence through motivation the way in which the predictive signals are processed during decision-making scenarios (e.g. Baines, Ruz, Rao, Denison, & Nobre, 2011; Rushworth & Behrens, 2008). In the social realm, both personal identity and emotional expressions are crucial sources of information to predict the behavior of others and to act in accordance with them (Ekman & Friesen, 2015; Frijda, 1988; Frijda & Mesquita, 1994; Frith & Frith, 2007). In our experiment, predicting on the basis of the emotion from the faces is a strategy consistent with approach/avoidance tendencies linked to the valence of emotions (see for example, Fridlund, 2014). Emotions, useful tools during negotiations, naturally lead to fast inferences about the internal states and dispositions of others (Fridlund, 2014). In accordance with this, the perceiver would focus attentional resources only on the signals with a higher predictive value of the future behavior of the agents (Fridlund, 2014). In this line, angry expressions act as a message of likely negative intentions, which would reduce the number of bets during the game, while expressed happiness leads to infer future positive behaviors, which would increase bets during the game. Similarly, in those situations where the identity was the predictive signal, results fit with the overgeneralization hypothesis (Todorov et al., 2005, 2008). Despite the decision of participants was guided by the relevant information at hand, when they were using the identity as a predictive cue of future outcomes, participants needed more time to respond when they were expecting gain but facing an angry person. Such slowing down of responses reflects conflict between identity expectations and emotions, processed in a non-volitional or automatic manner. This result is coherent with evidence of conflict in previous studies (Alguacil et al., 2015; Oosterhof & Todorov, 2008; Ruz et al., 2013; Ruz & Tudela, 2011). Lastly, it is worth to mention that although emotional content was early gathered, as discussed below, and did not suffer from any other source of conflict, participants needed more time to reach their decision during the Emotion compared to the Identity block. This slowing down has been previously observed in conflict paradigms (e.g. Egner, Etkin, Gale, & Hirsch, 2008) and in emotional categorization studies (Atkinson et al., 2005). One of the plausible explanations is that emotional expressions possess rich meaningfulness, so they would require a kind of processing more extended in time, which would lead to increased response latencies (see Egner and colleagues, 2008; Ben-David, Chajut, & Algom, 2012; Mogg, Holmes, Garner, & Bradley, 2008).

The P100 and the N1 potentials have been associated to perceptual processing and attentional orienting (Luck, 2005; Mangun & Hillyard, 1991). At this early level, in our experiment these potentials reflected only emotional, not identity, sensitivity. The amplitude of the P100 revealed an hemispheric asymmetry for the perception of emotional salient information, in coherence with the hypothesis that the right hemisphere is dominant for emotional content (see for example, Hagemann, Hewig, Naumann, Seifert, & Bartussek, 2003, 2005; Killgore & Yurgelun-Todd, 2007). This effect, however, must be taken with caution as the result was only marginally significant (p=0.52). At the same time, the N1 was sensitive to the strategic use of cues. The amplitude of this potential increased with happy emotional expressions anticipating gain compared to angry ones, signaling loss. This enhancement only took place when emotions were the relevant predictors of trial outcome. Crucially, when personal identity was used as cue, the same emotional expressions being ignored did not modulate the N1, which links the observed effects to the strategic use of emotions to generate predictions. Note that, as explained in the Methods section, during the identity block only trials in which happiness was paired with an identity predicting gain and those in which anger appeared with an identity predicting loss were included in the ERPs, and thus the two blocks were fully equated in terms of stimuli distribution. Hence, predictive emotional information may receive an attentional benefit at this early processing level to facilitate its preferential decoding (Vuilleumier, 2005), at least when it is explicitly employed to guide decisions (Eimer & Holmes, 2002; Holmes, Vuilleumier, & Eimer, 2003; Pizzagalli, Koenig, Regard, & Lehmann, 1999). Some studies have found a 'negative bias' given that such fast processing tends to occur with negative information (Carretié, Hinojosa, Martín-Loeches, Mercado, & Tapia, 2004; Luo et al., 2010). However, our results for the N1 reflect an increased response to positive valence. This is also consistent with previous studies suggesting an initial rapid global analysis of salient emotional information prior to a complex and fine visual categorization (Pizzagalli, Koenig, Regard, & Lehmann, 1999; Vuilleumier & Pourtois, 2007). Crude emotional information extracted from faces at this early point in time, could be a 'header' to prepare areas for perceiving more detailed information in subsequent steps (Sugase, Yamane, Ueno, & Kawano, 1999; Yang et al., 2011). In coherence with this proposal, studies focused on the correlation between ERP and fMRI data have proposed that ERPs on the 100 ms time-window seem to be related to the activity of the occipital extrastriate areas (Pourtois et al., 2005; Yang et al., 2011). As proposed by Haxby and colleagues (Haxby & Gobbini, 2011; Haxby et al., 2000) inferior occipital regions send inputs to other areas of the Core System such as the FG, mFG, and STS (Haxby & Gobbini, 2011; Haxby et al., 2000; Hoffman & Haxby, 2000). Information about categorical facial dimensions, such as race or instructed trustworthiness, is sometimes reflected in these early potentials (Ito & Urland, 2003; Ruz et al., 2013; Tortosa, Lupiáñez, et al., 2013). The absence of similar effects for the identity cue may be partially due to the lack of differences at low-level physical characteristic among facial identities (in contrast to Ito & Urland, 2003; Tortosa, Lupiáñez, et al., 2013) as they were equally distributed within all experimental conditions (they had the same size, location, contrast, spatial frequency and viewpoint). Also, the present study did not employ any anticipatory cue that allowed participants to bias their attention during the preparation stage before face appearance (see Ruz et al., 2013), which together with the uniformity at low-level characteristics may explain the absence of differences linked to identities at this early stage.

One hundred and fifty milliseconds after face onset, a more detailed decoding of both invariant and changeable features seems to be reflected on the VPP and the N170 (Eimer, 2000; Jeffreys, 1989; Joyce & Rossion, 2005; Rossion et al., 2000). Whereas the strategic use of emotion as predictive cue exerted influence on the amplitude of the VPP, identity did not. Its amplitude was heightened for angry faces that signaled likely loss. On the other hand, the effect of emotional conflict during the identity block was reflected on the N170, but not on the VPP. Identities displaying an expression of anger but linked to a likely gain significantly enlarged the N170. Although the VPP has been defined as the dipole counterpart of the N170 with equivalent functions and similar neural sources (Joyce & Rossion, 2005), our results do not completely support such idea (see also Bötzel, Schulze, & Stodieck, 1995; Wong et al., 2009). While some studies have proposed the inferior occipital and posterior fusiform gyri as neural generators of the VPP and, jointly with the STS, preferentially of the N170 potential (Bötzel et al., 1995; Caharel et al., 2002; Iidaka, Matsumoto, Haneda, Okada, & Sadato, 2006; Nguyen & Cunnington, 2014; Rossion et al., 2003), others have added the amygdala and the orbitofrontal cortex, both involved in emotion recognition (Adolphs, 2002; Wong et al., 2009), as the principal neural sources involved in the generation of the VPP. In addition, some fMRI studies, have found a functional coupling between the FG and frontal lobe areas during emotional conflict occurrence (Egner et al., 2008; Egner & Hirsch, 2005). All this evidence supports the VPP bias for emotional information and the influence of both emotion and identity on the N170 potential amplitude. With respect to the emotional conflict effect that we found at this stage, previous studies have found coincident effects where the valence congruency between words and emotional expression interact at the N170 (Zhu et al., 2010; see also Hinojosa et al., 2015). The current study adds a social context by studying conflict stemming from the incongruence between expectations arisen from two different facial aspects crucial for social adaptive communication, emotion and person identity. Thus, our study is the first to show an effect of emotional conflict at the level of the N170 potential generated by the incongruence between expectations related to emotion and those extracted from identity. Interestingly, this effect displays the same direction as behavioral data, where we only observed differences for identities predicting gains. On this line, our results suggest that the construction of the perceptual representation of faces when their identity is being used to predict a future gain is influenced by their ignored emotional valence. This supports the interactive nature of identity and facial emotional information from an early stage of processing (Eimer, 2011; Hinojosa et al., 2015; Jeffreys, 1989; Joyce & Rossion, 2005; Pitcher et al., 2007; Rossion et al., 2000; Vuilleumier & Pourtois, 2007).

The N2 potential appeared shortly after. Although previous studies have reported N2 modulations by both identity and emotional expressions (Ruz et al., 2013; Tortosa, Lupiáñez, et al., 2013), in our case this potential reflects the first stage of processing affected by expectations arising from facial identity. Those faces linked to probable future loss heightened N2 amplitudes. On this line, the N2 deflection has been related to the presentation of task-relevant stimuli and the anticipation of probable outcomes before a feedback signal (Baker & Holroyd, 2011). In the field of decision-making, the ACC, one of the neural sources of this potential (van

Veen & Carter, 2002), is associated to the encoding of the value (i.e. gain or loss) of an action before the decision is made (e.g. Rushworth & Behrens, 2008). Given the nature of our task demands, this effect on the N2 potential could be driven by a detection of the incongruence between the expectations associated to identity, future loss, and the motivation to achieve gains during the game. The lack of an effect of emotional conflict on this potential, related to cognitive control and conflict processing in classical studies (Folstein & Van Petten, 2008; Heil et al., 2000; Kopp et al., 1996), may be partially due to the nature of the stimuli employed. While the N2 potential seems to be influenced by conflict in stimulus-response classic paradigms (Folstein & Van Petten, 2008; Kopp et al., 1996) and also by emotional background information (Kanske & Kotz, 2010a, 2010b, 2011), it does not seem to be a suitable index for this effect in paradigms where face processing plays a central role (Ruz et al., 2013; Tortosa, Lupiáñez, et al., 2013).

The last potential in which we observed modulations in terms of value predictions and conflict was the P3b, which has been associated to outcome evaluation, reward processing and inhibition during incongruent situations (Moser et al., 2014; Neuhaus et al., 2010; Polich, 2007; Valle-Inclán, 1996; Wu & Zhou, 2009). This potential showed a hemispheric asymmetry between the amplitude of both hemispheres; its amplitude was larger for the right than for the left hemisphere. Expectations linked to facial identity influenced the

P3b amplitude for the right hemisphere. This asymmetry has been reported previously and it has been attributed to the association of the P3b to the fronto-parietal right hemisphere attentional network (Polich, 2007). At left locations, conflict was reflected as an enhancement by the emotion expressed by identities linked to likely gain. This parallels a similar result from a previous study where the authors found heightened amplitude for positive personal descriptions linked to partners in a decision-making paradigm (Moser et al., 2014). In our experiment, the identities linked to likely gain are conceptually similar to trustworthy partners as far as both the positive personal description and the association to probable gain to a particular identity is a signal of the most beneficial economic outcome. Under the rationale of the inhibition hypothesis, when the identity and the facial expression signaled likely gain (positive outcome) the implementation of inhibition mechanism was not necessary, but when both features anticipated contradictory outcomes (i.e. the identity predicted gain and the angry facial expression led to an expectation of loss) the implementation of inhibition was more necessary, which reduced the P3b amplitude (see Neuhaus et al., 2010; Valle-Inclán, 1996). Our results provide further evidence to support the influence on the P3b potential amplitude of identities and facial expressions (Campanella et al., 2000, 2013, 2002). It is worth noting that to accurately employ the identity as a cue in our experiments, participants first had to retrieve from memory the expectation of positive or negative outcomes linked to each face learned at the beginning of the block. The capture of the attention by non-relevant emotional content of the face, could have boosted specific memory operations, which in turn would trigger inhibition mechanisms to avoid unnecessary information (Polich, 2007). The lack of effects at the level of the P3b potential for the emotion as predictive cue may be partially caused by the slowing-down of responses during the emotion block, which may have delayed the emotional effect on the P3b deflection.

There are, however, some limitations in the present study that warrant further investigations. We only found the emotional conflict effect, at the behavioral and ERP levels, for those identities predicting gain. This result is in agreement with previous results by employing a similar paradigm (Alguacil et al., 2015), although the reason for this is still unknown, and it is not always replicated (but see Alguacil, Díaz, Kotz, Mestres-Missé, Tudela & Ruz, *in preparation*). In addition, increasing the response interval after viewing the face could be helpful to check the influence of emotion at the P3b stage. Future studies could also include an 'identity conflict' version to fully disentangle the interaction between invariant and changeable facial aspects. Finally, a complementary fMRI study would help to better understand the brain regions and functional connections involved in this type of conflict.

In conclusion, our results support an earlier capture of attentional resources (P1/N1) by emotional relevant information than by facial identities when their strategic use for outcome predictions is contrasted in a decision-making paradigm. Shortly after (VPP/N170), both facial aspects are processed in detail and interact by generating an emotional conflict effect when the expectations arising from expression and identity contradicted each other (in parallel with behavioral data). After this, expectations from facial identity exerted influence on processing stages (N200) linked to the evaluation of the value of actions to guide future decisions. Finally, the emotional conflict between both facial aspects was reflected at the level of decision-making process (P3b). At this level, conflictive items were inhibited in favor of the selection of those with the highest motivational value for predicting outcomes.

6.6. References

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, *12*(2), 169–177.

Alguacil, S., Tudela, P., & Ruz, M. (2013). Cognitive and affective control in a flanker word task: common and dissociable brain mechanisms. *Neuropsychologia*, *51*(9), 1663–1672. http://doi.org/10.1016/j.neuropsychologia.2013.05.020

Alguacil, S., Tudela, P., & Ruz, M. (2015). Ignoring facial emotion expressions does not eliminate their influence on cooperation decisions. Psicológica, 36, 309-335.

Atkinson, A. P., Tipples, J., Burt, D. M., & Young, A. W. (2005). Asymmetric interference between sex and emotion in face perception. *Perception & Psychophysics*, *67*(7), 1199–1213. http://doi.org/10.3758/BF03193553

Baines, S., Ruz, M., Rao, A., Denison, R., & Nobre, A. C. (2011). Modulation of neural activity by motivational and spatial biases. *Neuropsychologia*, *49*(9), 2489–2497. http://doi.org/10.1016/j.neuropsychologia.2011.04.029

Baker, T. E., & Holroyd, C. B. (2011). Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. *Biological Psychology*, *87*(1), 25–34. http://doi.org/10.1016/j.biopsycho.2011.01.010

Barrett, L. F. (2012). Emotions are real. *Emotion (Washington, D.C.)*, *12*(3), 413–429. http://doi.org/10.1037/a0027555

Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a Psychological Primitive. *Advances in Experimental Social Psychology*, *41*, 167–218. http://doi.org/10.1016/S0065-2601(08)00404-8

Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Brain Research. Cognitive Brain Research*, *17*(3), 613–620.

Baudouin, J. Y., Gilibert, D., Sansone, S., & Tiberghien, G. (2000). When the smile is a cue to familiarity. *Memory (Hove, England)*, *8*(5), 285–292. http://doi.org/10.1080/09658210050117717

Ben-David, B. M., Chajut, E., & Algom, D. (2012). The Pale Shades of Emotion: A Signal Detection Theory Analysis of the Emotional Stroop Task. *Psychology*, *03*(07), 537–541. http://doi.org/10.4236/psych.2012.37079

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *Journal of Cognitive Neuroscience*, *8*(6), 551–565. http://doi.org/10.1162/jocn.1996.8.6.551

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, *10*(1), 122–142. http://doi.org/10.1006/game.1995.1027

Bötzel, K., Schulze, S., & Stodieck, S. R. G. (1995). Scalp topography and analysis of intracranial sources of face-evoked potentials. *Experimental Brain Research*, *104*(1), 135–143. http://doi.org/10.1007/BF00229863

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305–327. http://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Caharel, S., Poiroux, S., Bernard, C., Thibaut, F., Lalonde, R., & Rebai, M. (2002). ERPs associated with familiarity and degree of familiarity during face recognition. *The International Journal of Neuroscience*, *112*(12), 1499–1512.

Campanella, S., Bourguignon, M., Peigneux, P., Metens, T., Nouali, M., Goldman, S., … De Tiège, X. (2013). BOLD response to deviant face detection informed by P300 event-related potential parameters: A simultaneous ERP–fMRI study. *NeuroImage*, *71*, 92–103. http://doi.org/10.1016/j.neuroimage.2012.12.077

Campanella, S., Hanoteau, C., Dépy, D., Rossion, B., Bruyer, R., Crommelinck, M., & Guérit, J. M. (2000). Right N170 modulation in a face discrimination task: an account for categorical perception of familiar faces. *Psychophysiology*, *37*(6), 796–806.

Campanella, S., Quinet, P., Bruyer, R., Crommelinck, M., & Guerit, J.-M. (2002). Categorical perception of happiness and fear facial expressions: an ERP study. *Journal of Cognitive Neuroscience*, *14*(2), 210–227. http://doi.org/10.1162/089892902317236858

Carretié, L., Hinojosa, J. A., Martín-Loeches, M., Mercado, F., & Tapia, M. (2004). Automatic attention to emotional stimuli: neural correlates. *Human Brain Mapping*, *22*(4), 290–299. http://doi.org/10.1002/hbm.20037

de Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of*

Cognitive Neuroscience, *14*(2), 199–209. http://doi.org/10.1162/089892902317236849

Dennis, T. A., Malone, M. M., & Chen, C.-C. (2009). Emotional face processing and emotion regulation in children: an ERP study. *Developmental Neuropsychology*, *34*(1), 85–102. http://doi.org/10.1080/87565640802564887

Derks, B., Stedehouder, J., & Ito, T. A. (2015). Social identity modifies face perception: an ERP study of social categorization. *Social Cognitive and Affective Neuroscience*, *10*(5), 672–679. http://doi.org/10.1093/scan/nsu107

Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2008). Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(6), 1475–1484. http://doi.org/10.1093/cercor/bhm179

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, *8*(12), 1784–1790. http://doi.org/10.1038/nn1594

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport*, *11*(10), 2319–2324.

Eimer, M. (2011). The Face-Sensitivity of the N170 Component. *Frontiers in Human Neuroscience*, *5*. http://doi.org/10.3389/fnhum.2011.00119

Eimer, M., & Holmes, A. (2002). An ERP study on the time course of emotional face processing. *Neuroreport*, *13*(4), 427–431.

Ekman, P., & Friesen, W. V. (2015). *Unmasking the Face: A Guide to Recognizing Emotions From Facial Expressions*. Malor Books.

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. http://doi.org/10.1111/j.1469-8986.2007.00602.x

Fridlund, A. J. (2014). *Human Facial Expression: An Evolutionary View*. Academic Press.

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*(5), 349–358. http://doi.org/10.1037/0003-066X.43.5.349

Frijda, N. H., & Mesquita, B. (1994). The social roles and functions of emotions. In S. Kitayama & H. R. Markus (Eds.), *Emotion and culture: Empirical studies of mutual influence* (pp. 51–87). Washington, DC, US: American Psychological Association.

Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology: CB*, *17*(16), R724–732. http://doi.org/10.1016/j.cub.2007.05.068

Gobbini, M. I., & Haxby, J. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, *45*(1), 32–41. http://doi.org/10.1016/j.neuropsychologia.2006.04.015

Hagemann, D., Hewig, J., Naumann, E., Seifert, J., & Bartussek, D. (2003). Physiological responses while watching emotional films: A multi-level approach. In *Journal of Psychophysiology* (Vol. 17, pp. 151–151). Hogrefe & Huber Publishers Rohnsweg 25, D-37085 Gottingen, Germany.

Hagemann, D., Hewig, J., Naumann, E., Seifert, J., & Bartussek, D. (2005). Resting brain asymmetry and affective reactivity: Aggregated data support the right-hemisphere hypothesis. *Journal of Individual Differences*, *26*(3), 139–154.

Haxby, J., & Gobbini, M. I. (2011). Distributed Neural Systems for Face Perception. In A. Calder, G. Rhodes, M. Johnson, & J. V. Haxby (Eds.), *Oxford Handbook of Face Perception* (pp. 93–110). OUP Oxford.

Haxby, J., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. http://doi.org/10.1016/S1364-6613(00)01482-0

Heil, M., Osman, A., Wiegelmann, J., Rolke, B., & Hennighausen, E. (2000). N200 in the Eriksen-Task: Inhibitory Executive Processes? *Journal of Psychophysiology*, *14*(4), 218–225. http://doi.org/10.1027//0269-8803.14.4.218

Hilimire, M. R., Mienaltowski, A., Blanchard-Fields, F., & Corballis, P. M. (2014). Age-related differences in event-related potentials for early visual processing of emotional faces. *Social Cognitive and Affective Neuroscience*, *9*(7), 969–976. http://doi.org/10.1093/scan/nst071

Hinojosa, J. A., Mercado, F., & Carretié, L. (2015). N170 sensitivity to facial expression: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *55*, 498–509. http://doi.org/10.1016/j.neubiorev.2015.06.002

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*(1), 80–84. http://doi.org/10.1038/71152

Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. *Brain Research. Cognitive Brain Research*, *16*(2), 174–184.

Iidaka, T., Matsumoto, A., Haneda, K., Okada, T., & Sadato, N. (2006). Hemodynamic and electrophysiological relationship involved in human face processing: evidence from a combined fMRI-ERP study. *Brain and Cognition*, *60*(2), 176–186. http://doi.org/10.1016/j.bandc.2005.11.004

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, *85*(4), 616–626. http://doi.org/10.1037/0022-3514.85.4.616

Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Experimental Brain Research*, *78*(1), 193–202.

Joyce, C., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. *Clinical Neurophysiology: Official Journal of the*

International Federation of Clinical Neurophysiology, *116*(11), 2613–2631. http://doi.org/10.1016/j.clinph.2005.07.005

Kanske, P., & Kotz, S. A. (2010a). Emotion Speeds up Conflict Resolution: A New Role for the Ventral Anterior Cingulate Cortex? *Cerebral Cortex*, bhq157. http://doi.org/10.1093/cercor/bhq157

Kanske, P., & Kotz, S. A. (2010b). Modulation of early conflict processing: N200 responses to emotional words in a flanker task. *Neuropsychologia*, *48*(12), 3661–3664. http://doi.org/10.1016/j.neuropsychologia.2010.07.021

Kanske, P., & Kotz, S. A. (2011). Positive emotion speeds up conflict processing: ERP responses in an auditory Simon task. *Biological Psychology*, *87*(1), 122–127. http://doi.org/10.1016/j.biopsycho.2011.02.018

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *17*(11), 4302–4311.

Killgore, W. D. S., & Yurgelun-Todd, D. A. (2007). The right-hemisphere and valence hypotheses: could they both be right (and sometimes left)? *Social Cognitive and Affective Neuroscience*, *2*(3), 240–250. http://doi.org/10.1093/scan/nsm020

Kopp, B., Rist, F., & Mattler, U. (1996). N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology*, *33*(3), 282–294.

Luck, S. J. (2005). *An Introduction to the Event-related Potential Technique*. MIT Press.

Luo, W., Feng, W., He, W., Wang, N.-Y., & Luo, Y.-J. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *NeuroImage*, *49*(2), 1857–1867. http://doi.org/10.1016/j.neuroimage.2009.09.018

Mangun, G. R., & Hillyard, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology. Human Perception and Performance*, *17*(4), 1057–1074.

Mogg, K., Holmes, A., Garner, M., & Bradley, B. P. (2008). Effects of threat cues on attentional shifting, disengagement and response slowing in anxious individuals. *Behaviour Research and Therapy*, *46*(5), 656–667. http://doi.org/10.1016/j.brat.2008.02.011

Moser, A., Gaertig, C., & Ruz, M. (2014). Social information and personal interests modulate neural activity during economic decision-making. *Frontiers in Human Neuroscience*, *8,* 31. http://doi.org/10.3389/fnhum.2014.00031

Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant and Child Development*, *10*(1-2), 3–18. http://doi.org/10.1002/icd.239

Neuhaus, A. H., Urbanek, C., Opgen-Rhein, C., Hahn, E., Ta, T. M. T., Koehler, S., … Dettling, M. (2010). Event-related potentials associated with Attention Network Test. *International Journal of Psychophysiology*, *76*(2), 72–79. http://doi.org/10.1016/j.ijpsycho.2010.02.005

Nguyen, V. T., & Cunnington, R. (2014). The superior temporal sulcus and the N170 during face processing: single trial analysis of concurrent EEG-fMRI. *NeuroImage*, *86*, 492–502. http://doi.org/10.1016/j.neuroimage.2013.10.047

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. http://doi.org/10.1073/pnas.0805664105

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia*, *45*(1), 75–92. http://doi.org/10.1016/j.neuropsychologia.2006.04.025

Parkinson, B. (1996). Emotions are social. *British Journal of Psychology (London, England: 1953)*, *87 (Pt 4)*, 663–683.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*(2), 184–187.

Pitcher, D., Walsh, V., & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Experimental Brain Research*, *209*(4), 481–493. http://doi.org/10.1007/s00221-011-2579-1

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS Evidence for the Involvement of the Right Occipital Face Area in Early Face Processing. *Current Biology*, *17*(18), 1568–1573. http://doi.org/10.1016/j.cub.2007.07.063

Pizzagalli, D., Koenig, T., Regard, M., & Lehmann, D. (1999). Affective attitudes to face images associated with intracerebral EEG source location before face viewing. *Brain Research. Cognitive Brain Research*, *7*(3), 371–7. http://doi.org/10.1016/S0926-6410(98)00040-8

Pizzagalli, D., Lehmann, D., Hendrick, A. M., Regard, M., Pascual-Marqui, R. D., & Davidson, R. J. (2002). Affective Judgments of Faces Modulate Early Activity (~160 ms) within the Fusiform Gyri. *NeuroImage*, *16*(3, Part A), 663–677. http://doi.org/10.1006/nimg.2002.1126

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148. http://doi.org/10.1016/j.clinph.2007.04.019

Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., & Vuilleumier, P. (2005). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: time course and topographic evoked-potentials mapping. *Human Brain Mapping*, *26*(1), 65–79. http://doi.org/10.1002/hbm.20130

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral Cortex (New York, N.Y.: 1991)*, *14*(6), 619–633. http://doi.org/10.1093/cercor/bhh023

Rossion, B., & Caharel, S. (2011). ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception. *Vision Research*, *51*(12), 1297–1311. http://doi.org/10.1016/j.visres.2011.04.003

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., & Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *Neuroreport*, *11*(1), 69–74.

Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *NeuroImage*, *20*(3), 1609–1624.

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*(4), 389–397. http://doi.org/10.1038/nn2066

Ruz, M., Madrid, E., & Tudela, P. (2013). Interactions between perceived emotions and executive attention in an interpersonal game. *Social*

Cognitive and Affective Neuroscience, 8(7), 838–844. http://doi.org/10.1093/scan/nss080

Ruz, M., & Tudela, P. (2011). Emotional conflict in interpersonal interactions. *NeuroImage, 54*(2), 1685–1691. http://doi.org/10.1016/j.neuroimage.2010.08.039

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*. Psychology Software Incorporated.

Schweinberger, S. R., & Soukup, G. R. (1998). Asymmetric relationships among perceptions of facial identity, emotion, and facial speech. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(6), 1748–1765.

Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, *400*(6747), 869–873. http://doi.org/10.1038/23703

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science (New York, N.Y.), 308*(5728), 1623–1626. http://doi.org/10.1126/science.1110589

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*(12), 455–460. http://doi.org/10.1016/j.tics.2008.10.001

Todorov, A., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*, *83*(5), 1051–1065. http://doi.org/10.1037/0022-3514.83.5.1051

Tooby, J., & Cosmides, L. (1990). The past explains the present. *Ethology and Sociobiology*, *11*(4), 375–424. http://doi.org/10.1016/0162-3095(90)90017-Z

Tortosa, M. I., Lupiáñez, J., & Ruz, M. (2013). Race, emotion and trust: an ERP study. *Brain Research*, *1494*, 44–55. http://doi.org/10.1016/j.brainres.2012.11.037

Tortosa, M. I., Strizhko, T., Capizzi, M., & Ruz, M. (2013). Interpersonal effects of emotion in a multi-round Trust Game.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., … Nelson, C. (2009). The NimStim set of facial expressions:

Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. http://doi.org/10.1016/j.psychres.2008.05.006

Tucker, D. M. (1994, March 8). Head sensor positioning network.

Valle-Inclán, F. (1996). The locus of interference in the Simon effect: an ERP study. *Biological Psychology*, *43*(2), 147–162. http://doi.org/10.1016/0301-0511(95)05181-3

van Veen, V., & Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology & Behavior*, *77*(4–5), 477–482. http://doi.org/10.1016/S0031-9384(02)00930-7

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594. http://doi.org/10.1016/j.tics.2005.10.011

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia*, *45*(1), 174–194.

Wager, T. D., Phan, K. L., Liberzon, I., & Taylor, S. F. (2003). Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. *NeuroImage*, *19*(3), 513–531.

Willis, J., & Todorov, A. (2006). First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science*, *17*(7), 592–598. http://doi.org/10.1111/j.1467-9280.2006.01750.x

Wong, T. K. W., Fung, P. C. W., McAlonan, G. M., & Chua, S. E. (2009). Spatiotemporal dipole source localization of face processing ERPs in adolescents: a preliminary study. *Behavioral and Brain Functions: BBF*, *5*, 16. http://doi.org/10.1186/1744-9081-5-16

Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, *1286*, 114–122. http://doi.org/10.1016/j.brainres.2009.06.032

Yang, Y., Gu, G., Guo, H., & Qiu, Y.-H. (2011). Early event-related potential components in face perception reflect the sequential neural activities. *Sheng Li Xue Bao: [Acta Physiologica Sinica]*, *63*(2), 97–105.

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *24*(28), 6258–6264. http://doi.org/10.1523/JNEUROSCI.4537-03.2004

Zhu, X., Zhang, H., Wu, T., Luo, W., & Luo, Y. (2010). Emotional conflict occurs at an early stage: evidence from the emotional face-word

Stroop task. *Neuroscience Letters*, *478*(1), 1–4. http://doi.org/10.1016/j.neulet.2010.04.036

Capítulo 7: Discusión General

La sección de Discusión General está organizada en cuatro partes. En la primera de ellas se presentan, de forma breve, los principales resultados obtenidos de cada serie. Los dos puntos siguientes se centran en la integración de los efectos de conflicto durante las interacciones personales, en el marco del control cognitivo y la neurociencia social y afectiva. En la última parte se presentan las conclusiones principales extraídas del presente trabajo.

7.1. Resumen general de resultados

A lo largo de estudios que componen esta tesis, nuestras investigaciones han ofrecido indicadores de conflicto cognitivo a nivel comportamental, electrofisiológico y de neuroimagen, generado por información neutra (Serie I) y emocional (Series I-IV), y en contextos simplificados de interacción social (Series II-IV).

Los principales resultados comportamentales obtenidos en la primera serie apoyan la adecuación del paradigma de flancos con palabras neutras y emocionales para permitir la comparación de la interferencia cognitiva y emocional. En este sentido, el conflicto generado por los esquemas de acción de los estímulos flanco, con independencia de su saliencia afectiva, es equivalente a nivel comportamental. Los resultados electrofisiológicos que obtuvimos muestran un mecanismo común a ambos tipo de conflicto en etapas tempranas, tradicionalmente relacionadas con el análisis perceptivo de la información. Este mecanismo común, en relación a las teorías clásicas de control, indica la existencia de una facilitación a nivel perceptivo durante las situaciones de conflicto en función del contexto de congruencia dado por el estímulo previo. Dicha facilitación, distinta entre los potenciales P1 y N170, parece corresponder a una mejora en la detección del conflicto en el primer caso y de la puesta en marcha de ajustes de control en el segundo. El descarte de una respuesta automática inapropiada y la selección de un esquema de acción válido en cada tarea tiene lugar en distintas etapas de procesamiento. Estos ajustes quedan reflejados en el potencial cortical N2 y la primera fase del P3 para la tarea de tipo cognitivo, y en la tarea de tipo afectivo recaen fundamentalmente en la última fase del P3. Nuestros resultados son coherentes con los hallazgos previos en la literatura para el potencial N2, sensible a la detección de respuestas dominantes, bien aprendidas, pero inadecuadas para la resolución de la tarea antes de que finalmente sean proporcionadas (Kopp, Rist, & Mattler, 1996; van Veen & Carter, 2002). Una vez en la fase de respuesta, donde se ha de indicar la decisión final, el coste en recursos cognitivos consecuencia de la inhibición queda reflejado en el potencial cortical P3, donde alcanzan mayor amplitud las respuestas de tipo congruente (e.g. Polich, 2007). En el caso de la tarea de tipo afectivo, la saliencia de la información emocional impacta sobre la etapa de respuesta, donde encontramos mayor amplitud para los ensayos de tipo incongruente. Esta última fase del potencial P3 es sensible al valor afectivo de la información presentada. A este nivel, la categorización de las respuestas asociadas a las situaciones de incongruencia en la tarea afectiva genera una mayor amplitud, dada su relevancia en términos motivacionales.

Los resultados de la Serie II muestran, en primer lugar, la viabilidad de las tareas de tipo económico adaptadas para el estudio de los efectos de conflicto en contextos de tipo social. Al mismo tiempo, son índice de la importancia del fenómeno del conflicto en el estudio de los mecanismos de control, ya que destacan la presencia de este efecto en multitud de contextos distintos, en este caso también incluyendo paradigmas que adaptan juegos económicos clásicos. En este ámbito, nuestros resultados indican que las expectativas generadas por la información personal guían, por encima de la influencia de otros factores, las decisiones tomadas por los participantes durante la tarea. Sin embargo, durante las situaciones en las que las expectativas asociadas de manera natural a la emociones, irrelevantes durante el juego, y las atribuidas a la identidad de los compañeros de juego, mediante instrucción, conllevan cursos de acción contrapuestos, la interferencia entre ambas queda reflejada en un coste atencional. Los participantes necesitan más tiempo para seleccionar su respuesta en dichas situaciones. Por tanto, a pesar de la irrelevancia de la información emocional durante la tarea, incluso cuando se proporcionan instrucciones explícitas acerca de su nulo peso para predecir las consecuencias futuras, esta genera una respuesta automática que debe ser regulada mediante estrategias de control.

Durante la Serie III, replicamos los efectos de conflicto encontrados en la serie experimental previa, siendo estos extensibles, a diferencia de los resultados anteriores, a las identidades asociadas a expectativas de pérdida. Los resultados de neuroimagen indican la participación de estructuras cerebrales asociadas a la atribución de recompensa a los resultados esperables de determinadas acciones durante el procesamiento de la identidad como clave predictiva. Estas asociaciones permiten generar un mapa de expectativas de cada identidad para facilitar la selección de aquellas con valor de recompensa más alto. Cuando estas expectativas son elemento suficiente para seleccionar la respuesta más adaptativa, conforme a las exigencias de la tarea, al tiempo que las expectativas naturalmente asociadas a las expresiones emocionales no generan una competición por estar solapadas completamente con las primeras, intervienen regiones participantes en el procesamiento de los rasgos invariantes del rostro (i.e. la identidad). Sin embargo, cuando las expectativas de identidad y emoción entran en competición, el conflicto es detectado por la CCA. Tras la señal de conflicto, la interferencia es solventada mediante la intervención de esta estructura en coordinación con regiones frontales, que potencian las vías de procesamiento asociadas a las respuestas correctas, y otras más posteromediales, probablemente relevantes para la focalización en la dimensión relevante del rostro (i.e. identidad) de los estímulos y la recuperación de las atribuciones de recompensa de esta almacenadas en memoria.

En la última serie del presente trabajo empleamos HDEEG para determinar el curso de procesamiento de la emoción y la identidad como claves predictivas en contextos sociales, con el fin añadido de determinar a qué niveles de procesamiento se produce la interferencia entre ambos cuando la identidad es la dimensión relevante de la tarea. En esta última serie encontramos una captura temprana de la atención por la emoción como clave (P1), que también posibilitó una discriminación más rápida de las distintas asociaciones de ganancia y pérdida (N1 y VPP). Las predicciones vinculadas a la identidad se mostraron significativas para los potenciales corticales N2, sensible a la atribución de valor a las acciones en función de sus consecuencias, y P3b, clásicamente presentado como un componente que refleja la inhibición de respuestas inapropiadas. Durante la tarea donde la identidad era el elemento predictivo, la influencia de las emociones fue significativa en una primera etapa asociada a la discriminación de la identidad (N170) así como en la etapa de selección de la respuesta adecuada (P3b).

7.2. Automaticidad de la información emocional: formación de esquemas de acción

La relevancia de la información emocional para responder de forma adaptativa a las demandas del entorno es un hecho establecido en múltiples dominios. Permiten, por ejemplo, dar una respuesta rápida y eficiente ante situaciones de peligro (Cohen, 2005; LeDoux, 2012). Más allá de impulsar una serie de respuestas automáticas ante eventos potencialmente amenazantes, juegan un papel central en contextos de tipo social (e.g. Parkinson, 1996). Entre las funciones asignadas a los fenómenos afectivos, una de las más destacadas es su papel como herramienta de comunicación durante las interacciones personales. Permiten transmitir de forma rápida y eficiente no solo la emoción en sí sino también creencias, disposición general e intenciones de otras personas (Keltner & Gross, 1999; Keltner & Haidt, 1999).

Como muestran nuestros resultados en su conjunto, la información afectiva puede generar interferencia bien por su valencia positiva o negativa (Serie I), o bien por el contenido informativo de la misma acerca de las predisposiciones de las personas con las que interactuamos, dando lugar a una categorización más compleja en relación al valor de las consecuencias predichas de sus acciones (Series II, III y IV). De aquí se concluye que los estímulos con saliencia emocional tienen la habilidad de captar nuestra atención e influir en el curso de nuestras acciones en multitud de contextos. Aunque los efectos de conflicto emocional son extrapolables a situaciones de tipo social, como indican nuestros datos, el modo en el que se inclina la balanza de activación/inhibición entre los distintos esquemas de acción disponibles puede diferir entre la tarea empleada en la Serie I y la empleada en las series subsiguientes.

Como introducíamos en el primer capítulo de la tesis, la selección de un mapa de acción que se ajuste a las demandas de la situación depende, en primer lugar, de la relación entre la activación e inhibición de los esquemas disponibles (Norman & Shallice, 1986). Debido a un mecanismo lateral de competición entre los esquemas que se han asociado a través del aprendizaje a determinadas situaciones, el que de ellos consiga alcanzar un mayor umbral de activación es el que prevalece sobre el resto (Norman & Shallice, 1986). La determinación del valor de activación de los esquemas vinculados a la tarea parece depender también de la interrelación existente entre ellos. Así, la activación de una representación específica en un sistema favorece la propagación de esta a otras estructuras, aumentando en consecuencia el valor de activación en los diferentes sistemas de procesamiento que intervengan (Norman & Shallice, 1986).

En nuestro caso, la competición entre esquemas viene proporcionada por la relación simbiótica entre los juicios sobre la confianza dados por la identidad, que contribuyen a determinar si hemos o no de actuar de manera cooperativa con ellos, y las expectativas que se generan en base a las emociones que muestran. La correlación entre estos juicios, basados en ambos rasgos faciales, es muy alta (e.g. Todorov, Said, Engell, & Oosterhof, 2008). Por una parte, las emociones generan en sí mismas tendencias de acercamiento y alejamiento (Wyer, 2014), dependiendo de las circunstancias. Los aspectos primitivos derivados de las emociones contribuyen, además, a la formación de impresiones como ocurre con los juicios acerca de la confianza que desprende una persona en particular. Si la valencia afectiva asignada a la emoción es positiva entonces se esperan una serie de actitudes positivas por parte de esa persona, entre ellas la de cooperar. Lo mismo ocurre en el extremo opuesto; una emoción negativa llevará a generar un juicio también de valor negativo en otras dimensiones basadas en el rostro. Esta relación estrecha entre ambos elementos indica cierto grado de dependencia entre los esquemas de acción que se activan cuando ambos son percibidos. En nuestro paradigma, mientras que la valoración de las emociones surge de manera automática, la prevalencia de los juicios centrados en la identidad es asignada mediante el aumento de la activación de las representaciones asociadas a través del mecanismo vertical ejercido por el SAS, con un peso mayor cuanto menor es la práctica adquirida durante la tarea. La interrelación entre representaciones puede derivar en que ambos elementos, identidad y emoción, acaben formando parte un mismo esquema general de acción, o que uno esté en cierto sentido incluido en el otro. En la vida diaria las predicciones derivadas de ambos rasgos, y en consecuencia la activación de los esquemas de acción asociados, son una forma simple de aproximarse a las situaciones de interacción personal. No solo de ambas se pueden extraer inferencias de una forma más o menos automática sino que además permiten, sin necesidad de invertir gran cantidad de recursos, solucionar determinados conflictos de forma eficiente y bastante similar. Por tanto, en un contexto como el que se da en los paradigmas de juego económico, ambos elementos tienen un alto grado de activación para generar predicciones, en muchos casos la coordinación de la información proporcionada por identidad y emoción puede llegar a ser inseparable. Estas coincidencias pueden generar conflicto cuando las expectativas van en direcciones opuestas pero también puede permitir que, de forma más flexible la información del rasgo irrelevante (i.e. emoción) se supedite a las limitaciones, guiadas mediante instrucción explícita, que antepone la característica relevante (i.e. la identidad). Por tanto, la dependencia entre identidad y emoción puede generar, hasta cierto punto, un ajuste mejor de las estrategias de control aplicadas sobre la emoción.

Si durante la formación de las impresiones acerca de otras personas, a lo largo del aprendizaje, la emoción es un elemento que contribuye en un alto grado, cabe esperar que su empleo en situaciones futuras para generar nuevas impresiones sea necesario o al menos relevante. La evidencia acumulada a favor no solo de la utilidad de ambos rasgos faciales para hacer predicciones sino también la de la emoción como elemento vertebrador de los juicios de confianza se suma, haciendo que ante situaciones similares estos esquemas se activen de manera conjunta y se complementen de algún modo. La imposibilidad de evitar la influencia de la expresión emocional en contextos de interacción puede ser explicada desde esta interrelación entre emoción e identidad. Asimismo, esto ayudaría a explicar por qué hay cierta variabilidad, a lo largo de los estudios, con respecto al efecto de la emoción sobre la tasa de cooperación.

7.3. Automaticidad de la emoción en la decodificación de la identidad: de la percepción a la acción

Como muestran los resultados obtenidos en la última serie de la tesis, la emoción, como clave predictiva del comportamiento de otros y de las consecuencias esperables de sus acciones, captura rápidamente la atención ya desde etapas perceptivas. Esta captura temprana está justificada no solo por su valor a la hora de propiciar respuestas adaptativas sino también por su especial implicación en la generación de atribuciones personales para guiar el comportamiento (Frijda, 1988; Frijda & Mesquita, 1994; Tooby & Cosmides, 1990). La asociación entre el valor positivo o negativo de las consecuencias esperables del comportamiento de otras personas en base a las emociones parece suceder antes que para la identidad. Sin embargo, cuando se atiende a su papel, a pesar de ser irrelevante, durante el empleo de la identidad como clave predictiva, las expresiones emocionales solo influyen en la discriminación de la identidad una vez esta última ha sido descodificada, como queda reflejado en los efectos sobre el potencial cortical N170 en la Serie IV.

Esta intrusión por parte de las expectativas asociadas a las emociones se produce una vez es posible su integración con las inferencias de valor proporcionadas por la identidad. Por tanto, aunque las emociones poseen la capacidad de captar los recursos atencionales desde etapas tempranas, su influencia está supeditada a la elaboración de asociaciones específicas entre cada identidad y su valor de predicción en términos de si las consecuencias esperables son positivas o negativas. Solo una vez ambos son tenidos en cuenta, tiene lugar un efecto de conflicto entre las expectativas que de ellos se desprenden. Esta interferencia parece no estar determinada por las características físicas atribuibles a cada rasgo facial sino más bien por el valor de recompensa asociado a cada uno de ellos. Este hecho se ve reforzado por los resultados obtenidos en la Serie III, donde vemos que para la resolución del conflicto generado por las expresiones faciales, se produce una coordinación funcional entre regiones cerebrales implicadas en la recuperación de las asociaciones en base a la recompensa esperable de cada elemento. Por tanto, es durante la fase de comparación de las distintas probabilidades de obtener un beneficio atendiendo a uno u otro de los dos elementos faciales, lo que determina la competición y no tanto las posibles inferencias tempranas durante la fase puramente perceptiva. Cabe la posibilidad de que cuando las expectativas extraídas de la identidad y la emoción sean congruentes, estas se beneficien mutuamente aumentando la activación de los esquemas de acción comunes a ambas así como su representación, lo que facilitaría la selección de una respuesta en base a la identidad. Como muestran los resultados obtenidos con RMf, la elección de una opción adecuada en el caso de las situaciones de congruencia no requiere de ajustes estratégicos ya que la discriminación de la identidad, con las atribuciones específicas que esta posee, es suficiente para responder. Esto queda reflejado en los resultados obtenidos en la Serie III, donde el giro fusiforme, principal fuente generadora del potencial N170, participa al mismo tiempo que otras áreas de la corteza extraestriada en la selección de una estrategia. Cuando se produce una situación de conflicto a nivel de expectativas, aunque este conjunto de regiones participe en su procesamiento como sugieren los resultados del potencial N170 en la Serie IV, la mera discriminación de la identidad no sirve para regular la repuesta. Otras regiones relacionadas con la detección de la posible reducción de la recompensa asociada a un ítem pueden jugar un papel más relevante a este respecto. Este es el caso de la CCA dorsal (e.g. Bush et al., 2002).

A diferencia de los hallazgos en literatura previa de una mayor participación de la porción rostral de la CCA en la detección de la ocurrencia de conflicto de tipo emocional y el incremento de la comunicación entre esta y otras regiones límbicas, como la amígdala (e.g. Egner et al., 2008; Etkin et al., 2006), nuestros resultados no muestran tal especificidad. En cambio, encontramos que la detección es realizada por una porción de la CCA considerada de dominio general, es decir, inespecífica del tipo de estímulo o de su saliencia emocional. Al tratarse de un efecto de conflicto a nivel de expectativas asociadas un probable beneficio, la monitorización de interferencia es realizada por regiones relacionadas con el procesamiento del valor de recompensa de los estímulos. Nuestros resultados son coherentes con la literatura que defiende que la CCA es necesaria en un amplio abanico de funciones cognitivas, más allá de la mera monitorización de conflicto, entre ellas el procesamiento de recompensa asociado a diferentes estímulos. La relación entre los resultados obtenidos en las series III y IV revela algunos aspectos críticos acerca de cómo el conflicto emocional se solventa durante las interacciones personales.

Como apuntan nuestros resultados en conjunto, la respuesta de la CCA estaría más relacionada con la detección de una posible disminución en la recompensa futura esperable en situaciones de interferencia entre identidad y emoción (ver Goldstein et al., 2008; Wei, Wang, & Ji, 2016). Los factores motivacionales de la recompensa parecen ejercer una influencia distinta en estos índices de la que se recoge típicamente en la literatura sobre los mecanismos de control (e.g. Baker & Holroyd, 2011). Una vez detectada esta reducción, las estrategias de regulación sobre la acción para elegir una respuesta dependerían de la comparación del valor de recompensa, en el contexto de la tarea, que poseen identidad y emoción. Para esto intervendrían regiones frontales asociadas a la inhibición de las respuestas inapropiadas y otras posteromediales, como el precúneo y la CC posterior, necesarias para la recuperación de los valores de recompensa asignados a cada estímulo, que habrían sido almacenados en memoria tras ser aprendidos en el curso de la tarea. Estos serían adquiridos gracias al refuerzo que obtienen ensayo a ensayo, momento en el que participan el núcleo caudado y el putamen. Esta etapa de inhibición y recuperación de valores de recompensa para, finalmente, seleccionar la respuesta más apropiada queda reflejado no solo en los resultados obtenidos mediante RMf sino además en los efectos de conflicto obtenidos en el potencial cortical P3b en la serie última del trabajo. En conjunto, esto podría explicar, aunque de manera especulativa, por qué no encontramos efectos de conflicto a nivel del potencial N2 (aunque tampoco los encontramos durante la Serie I para la tarea emocional) y sí en etapas posteriores como el P3b, sensible procesamiento de la recompensa esperable.

Aunque hemos obtenido resultados válidos para explicar cómo el conflicto emocional tiene lugar en contextos de tipo social y de qué modo los mecanismos de control intervienen en su resolución, hay algunas incógnitas que permanecen sin respuesta. Por un lado, en investigaciones futuras sería interesante ampliar el foco de análisis a otros fenómenos relacionados con el conflicto como son los efectos secuenciales de conflicto. Esto ayudaría a entender los ajustes de las estrategias de control en situaciones de alto o bajo conflicto. Finalmente, uno de los resultados más difíciles de encajar es la ausencia de efectos de conflicto en el potencial cortical N2. Un estudio más detallado al respecto, centrado acaso en su relación con el potencial P3, sensible al procesamiento de expectativas, ayudaría a arrojar algo de luz sobre esta cuestión.

7.4. Conclusions

› Emotional conflict can also be studied in social paradigms, which is crucial given the social nature of affective phenomena.

- Emotions lead to inferences about the probable behaviors of other people, in the context of economic interpersonal exchanges, in an automatic or mandatory way.
- Even when emotions are irrelevant for the achievement of goals during the task, they influence the processing of facial identity when the latter is predictive of the behavior of others.
- While emotions do not disrupt the choices based on identity, their influence is reflected on the time that people need to make decisions about the most appropriate economic investment.
- Such modulation first takes place once identity and emotion have been decoded, as reflected on the conflict effect reflected on the amplitude of the N170.
- The degree of overlap between emotion and identity representations in terms of their reward value also affects response stages, as shown by results at the level of the P3b.
- Situations where the expectations from both facial features overlap increase the participation of face-sensitive areas, such as the fusiform gyrus and other extrastriate regions, for the decoding of identity.
- Domain-general conflict-related areas (anterior cingulate cortex) and other frontal regions (superior frontal gyrus) are recruited for the

detection of emotional conflict and the implementation of control mechanisms during personal interactions,

› The anterior cingulate cortex functionally interacts with frontal and posteromedial areas related to the evaluation of the reward value of personal identities held in memory, to prevent the selection of inappropriate automatic responses.

7.5. Referencias

Baker, T. E., & Holroyd, C. B. (2011). Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. *Biological Psychology*, *87*(1), 25-34. http://doi.org/10.1016/j.biopsycho.2011.01.010

Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., & Rosen, B. R. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(1), 523-528. http://doi.org/10.1073/pnas.012470999

Cohen, J. D. (2005). The Vulcanization of the Human Brain: A Neural Perspective on Interactions Between Cognition and Emotion. *Journal of Economic Perspectives*, *19*(4), 3-24.

Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2008). Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(6), 1475-1484. http://doi.org/10.1093/cercor/bhm179

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron*, *51*(6), 871-882. http://doi.org/10.1016/j.neuron.2006.07.029

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*(5), 349-358. http://doi.org/10.1037/0003-066X.43.5.349

Frijda, N. H., & Mesquita, B. (1994). The social roles and functions of emotions. En S. Kitayama & H. R. Markus (Eds.), *Emotion and culture: Empirical studies of mutual influence* (pp. 51-87). Washington, DC, US: American Psychological Association.

Goldstein, R. Z., Parvaz, M. A., Maloney, T., Alia-Klein, N., Woicik, P. A., Telang, F., … Volkow, N. D. (2008). Compromised sensitivity to monetary reward in current cocaine users: an ERP study. *Psychophysiology*, *45*(5), 705-713. http://doi.org/10.1111/j.1469-8986.2008.00670.x

Keltner, D., & Gross, J. J. (1999). Functional Accounts of Emotions. *Cognition and Emotion*, *13*(5), 467-480. http://doi.org/10.1080/026999399379140

Keltner, D., & Haidt, J. (1999). Social Functions of Emotions at Four Levels of Analysis. *Cognition and Emotion*, *13*(5), 505-521. http://doi.org/10.1080/026999399379168

Kopp, B., Rist, F., & Mattler, U. (1996). N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology*, *33*(3), 282-294.

LeDoux, J. (2012). Rethinking the Emotional Brain. *Neuron*, *73*(4), 653-676. http://doi.org/10.1016/j.neuron.2012.02.004

Norman, D., & Shallice, T. (1986). Attention to Action: Willed and Automatic Control of Behavior. En R. Davidson, R. Schwartz, & D. Shapiro (Eds.), *Consciousness and Self-Regulation: Advances in Research and Theory IV*. Plenum Press.

Parkinson, B. (1996). Emotions are social. *British Journal of Psychology (London, England: 1953)*, *87 (Pt 4)*, 663-683.

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128-2148. http://doi.org/10.1016/j.clinph.2007.04.019

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*(12), 455-460. http://doi.org/10.1016/j.tics.2008.10.001

Tooby, J., & Cosmides, L. (1990). The past explains the present. *Ethology and Sociobiology*, *11*(4), 375-424. http://doi.org/10.1016/0162-3095(90)90017-Z

van Veen, V., & Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology & Behavior*, *77*(4–5), 477-482. http://doi.org/10.1016/S0031-9384(02)00930-7

Wei, P., Wang, D., & Ji, L. (2016). Reward expectation regulates brain responses to task-relevant and task-irrelevant emotional words: ERP evidence. *Social Cognitive and Affective Neuroscience*, *11*(2), 191-203. http://doi.org/10.1093/scan/nsv097

Wyer, R. S. (2014). *The Automaticity of Everyday Life: Advances in Social Cognition*. Psychology Press.

Capítulo 8: Abstract

Humans posses a remarkable ability to quickly and flexibly adapt their behavior to current circumstances. Environmental changes imply alterations in the state of affairs, and these include social interactions. In many situations, the information available to the senses lead to multifaceted representations of the same element. Such multiplicity may generate conflict when the representations result in opposite action schemes (e.g. Botvinick, Braver, Barch, Carter, & Cohen, 2001; Botvinick, Cohen, & Carter, 2004; Kornblum, Hasbroucq, & Osman, 1990). This issue has been long studied on the field of cognitive control. However, most of research to date has employed neutral non-social stimuli (although see Bush, Luu, & Posner, 2000). Furthermore, the vast majority of studies have not offered information about how conflict with affective material affects different processing stages.

Experimental series I of the present thesis was aimed at comparing affective and neutral stimuli in a conflict paradigm. For this, we used a flanker word task, modified from the one by Ochsner and cols. (Ochsner, Hughes, Robertson, Cooper, & Gabrieli, 2009). We recorded participant's brain activity by means of a high-density electroencephalography (HDEEG) system to explore how affective and cognitive conflict influenced perceptual and decision-related stages of information processing. Results from this first experiment revealed that both conflicts were equivalent in terms of their behavioral effects. However, as shown by electrophysiological data, they were mediated by common and dissociable neural mechanisms. Early processing stages reflected on the P1 and N170 potentials, showed parallel sequential conflict effects. Cognitive conflict heightened the amplitude of the N2 and the potentials, whereas affective conflict only modulated the amplitude of the last portion of the P3.

The following series in the thesis transferred the experimental situation to a context closer to social interrelations. We are social animals and thus most of our routines, thoughts and behaviors depend on our interactions with other people (Frith & Frith, 2007; Wyer, 2014). This is especially true for emotional phenomena, as they are strongly related to the social realm (Parkinson, 1996). Personal information provided by facial identity and emotional expressions are two useful clues for communication in social contexts (e.g. Keltner & Haidt, 1999). Both facial features offer information about the intentions, thoughts, beliefs, emotions and internal states of others (Frijda, 1988; Frijda & Mesquita, 1994). Sometimes, expectations about future consequences based on identity and emotional expression may generate incongruent inferences about what the other people are going to do the next. In these cases, a conflict between both facial elements may arise, which calls for control of the automatic but inappropriate responses for the situation. How we adjust our behavior to deal with conflict stemming from competitive social clues is an issue still far from being understood. Thus, along three different experimental series we explored emotional conflict in social contexts.

In Experimental Series II, III and IV, we employed an adapted Trust Game (Tortosa, Lupiáñez, & Ruz, 2013; Tortosa, Strizhko, Capizzi, & Ruz, 2013). In this paradigm participants have to make economic decisions based on personal information linked to the facial identity of the people with whom they interact. The game partner displayed non-predictive emotional expressions (i.e. happy vs. angry) that had to be ignored.

In Experimental Series II, we evaluated emotional conflict at the behavioral level generated by the incongruence between the expectations linked to personal identity and emotional expressions (i.e. a cooperative person looking angry or a non-cooperative partner displaying an expression of happiness). As expected, when both facial clues led to opposite expectations about the most probable behavior of the partners, we observed an effect of conflict on the time needed to make the choices.

Next, we aimed to explore the neural mechanisms underlying the detection and resolution of the emotional conflict during interpersonal interactions. For this purpose, in the Experimental Series III functional magnetic resonance images (fMRI) were obtained during the performance of the task described above. As shown by fMRI results, when participants were facing a cooperative partner versus when they interacted with a non-cooperative one, the putamen and the body of the caudate were recruited presumably because of their role in computing reward values linked to positive predicted consequences. When the decision about identity did not entail any conflict, the most significant neural response was localized at the fusiform gyrus, a face-sensitive region involved in the extraction of facial invariant features such as identity. Neuroimaging results also revealed a set of neural areas engaged during conflictive situations, including the anterior cingulate cortex and the superior frontal gyrus. The anterior cingulate cortex was functionally coupled with other areas such as the inferior frontal gyrus and the posterior cingulate cortex/precuneus during conflict. This functional communication may be related to the inhibition of automatic inadequate responses led by emotion intrusion.

Finally, at Experimental Series IV, we aimed at exploring the processing levels that reflected the strategic employment of either the identity or the emotion to predict interpersonal outcomes. In addition, we analyzed the stages at which the irrelevant emotional information influenced the decoding of facial identity. Electrophysiological results revealed that when emotion was a relevant source of predictive information early perceptive stages were enhanced, as reflected on the P1, N1 and VPP potentials. Using identity for the same predictions enhanced the amplitude of the N2 and the

P3b potentials. In addition, the influence of the emotional irrelevant content from the faces was reflected on the N170 and P3b potentials. In sum, although emotions were fast decoded, they did not exert their influence on the processing of identity until both features had been decoded, at the level of the N170 potential.

Conclusions

Results from the Experimental Series I suggest that the first levels of information processing, mainly perceptual analyses, deal with cognitive and affective conflicts through the same mechanisms. By contrast, later stages involved in cognitive control and response planning diverge according to the nature of the task requirements generating the conflict. As the results from Experimental Series II-III and IV show, when emotional conflict is studied in social contexts, social constructs and emotions lead to different expectations about the proximal behavior of others (Fischer & Manstead, 2008; Haidt, 2001; Keltner & Haidt, 1999; Oosterhof & Todorov, 2008). The inconsistency between personal predispositions of cooperation and irrelevant emotional facial information leads to opposite expectations and increases demands on decision-making, which are reflected in slower response times. As neuroimaging data revealed, decisions guided by facial identity are biased by the reward value (i.e. gain or loss) of action-outcomes linked to each identity.

Facial identities predicting gain strongly activated relevant areas from the dorsal striatum involved in the formation of stimulus-action-reward associations. When their decoding of was not hindered by irrelevant and non-predictive emotional information, then occipital and temporal areas from the distributed neural system for face processing, relevant for the configuration of the identity, implemented the most adequate set of actions (Gobbini & Haxby, 2007; Haxby & Gobbini, 2011; Haxby et al., 2000, 2002). However, incongruent situations, call for additional conflict detection and control implementation mechanisms operated by a network including middle-frontal, frontolateral and a posteromedial areas. Electrophysiological results support an earlier capture of attentional resources by emotional relevant information than by facial identities when their strategic use for outcome predictions is contrasted in a decision-making paradigm. Shortly after, both facial aspects are processed in detail and interact by generating an emotional conflict effect when the expectations arising from expression and identity contradicted each other (in parallel with behavioral data). After this, expectations from facial identity exerted influence on processing stages linked to the evaluation of the value of actions to guide future decisions. Finally, the emotional conflict between both facial aspects was reflected at the level of decision-making process. At this level, conflictive items were inhibited in favor of the selection of those with the highest motivational value for predicting outcomes.

In sum, our results show that explicitly ignored emotions influence responses in a mandatory manner during interpersonal interactions. Our study, for very first time, shows how control mechanisms operate when emotional conflict emerges from incongruent expectations and non-overlapping action tendencies associated to invariant and changeable facial aspects during personal interactions.

References

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624-652.

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. http://doi.org/10.1016/j.tics.2004.10.003

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*(6), 215-222.

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*(5), 349-358. http://doi.org/10.1037/0003-066X.43.5.349

Frijda, N. H., & Mesquita, B. (1994). The social roles and functions of emotions. En S. Kitayama & H. R. Markus (Eds.), *Emotion and culture: Empirical studies of mutual influence* (pp. 51-87). Washington, DC, US: American Psychological Association.

Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology: CB*, *17*(16), R724-732. http://doi.org/10.1016/j.cub.2007.05.068

Keltner, D., & Haidt, J. (1999). Social Functions of Emotions at Four Levels of Analysis. *Cognition and Emotion, 13*(5), 505-521. http://doi.org/10.1080/026999399379168

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility--a model and taxonomy. *Psychological Review*, *97*(2), 253-270.

Ochsner, K. N., Hughes, B., Robertson, E. R., Cooper, J. C., & Gabrieli, J. D. E. (2009). Neural systems supporting the control of affective and cognitive conflicts. *Journal of Cognitive Neuroscience*, *21*(9), 1842-1855. http://doi.org/10.1162/jocn.2009.21129

Parkinson, B. (1996). Emotions are social. *British Journal of Psychology (London, England: 1953), 87 (Pt 4)*, 663-683.

Tortosa, M. I., Lupiáñez, J., & Ruz, M. (2013). Race, emotion and trust: an ERP study. *Brain Research*, *1494*, 44-55. http://doi.org/10.1016/j.brainres.2012.11.037

Tortosa, M. I., Strizhko, T., Capizzi, M., & Ruz, M. (2013). Interpersonal effects of emotion in a multi-round Trust Game.

Wyer, R. S. (2014). *The Automaticity of Everyday Life: Advances in Social Cognition*. Psychology Press.