# Towards an ecological approach to understanding the neurophysiological bases of human error-monitoring

Gonçalo Padrão

**ADVERTIMENT**. La consulta d’aquesta tesi queda condicionada a l’acceptació de les següents condicions d'ús: La difusió d’aquesta tesi per mitjà del servei TDX (**www.tdx.cat**) i a través del Dipòsit Digital de la UB (**diposit.ub.edu**) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d’investigació i docència. No s’autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d’un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s’autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

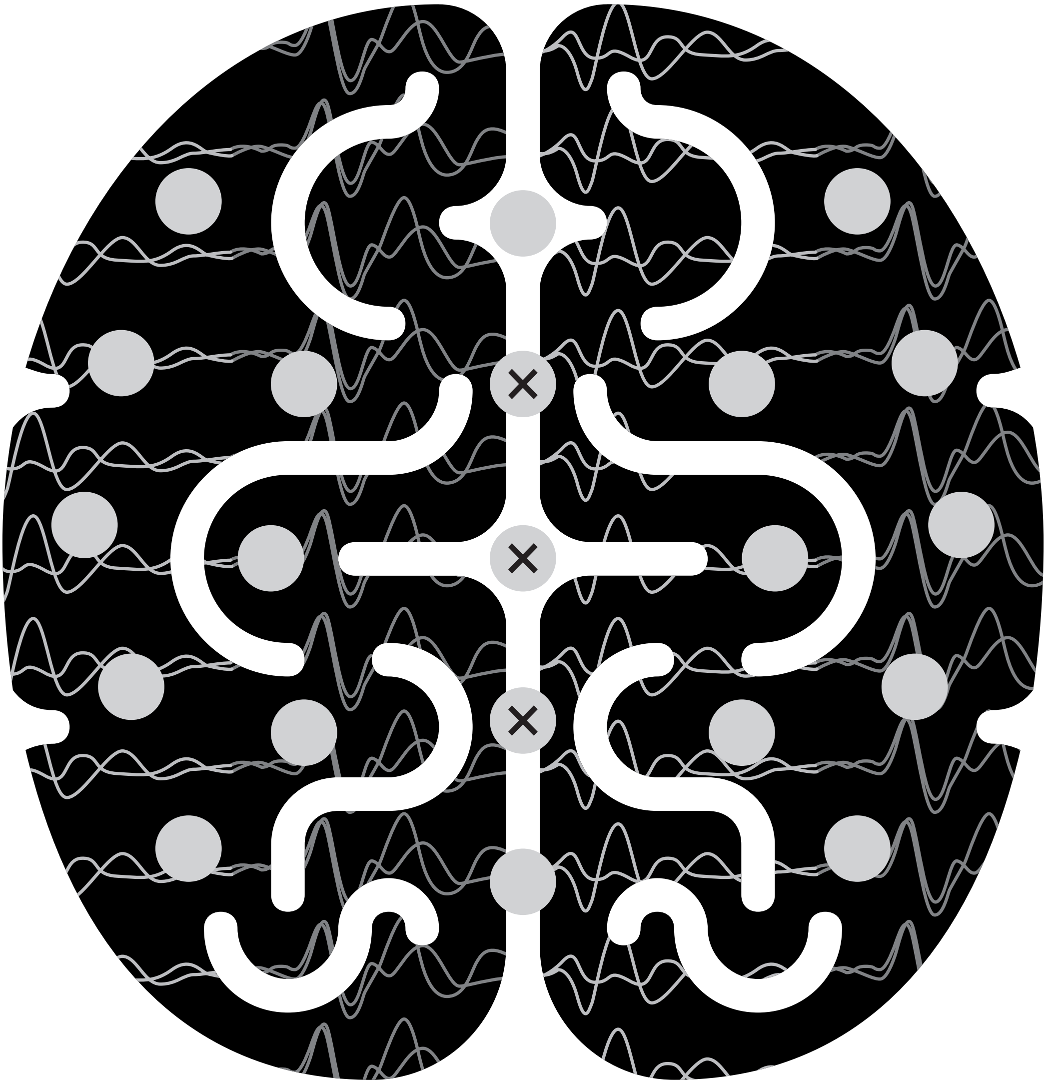
**ADVERTENCIA**. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (**www.tdx.cat**) y a través del Repositorio Digital de la UB (**diposit.ub.edu**) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**WARNING**. On having consulted this thesis you’re accepting the following use conditions: Spreading this thesis by the TDX (**www.tdx.cat**) service and by the UB Digital Repository (**diposit.ub.edu**) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it’s obliged to indicate the name of the author.

PhD Thesis

# Towards an ecological approach to understanding the neurophysiological bases of human error-monitoring

**Gonçalo Padrão**



**Doctoral Program in Biomedicine**

# Towards an ecological approach to understanding the neurophysiological bases of human error-monitoring

This thesis is presented by
**Gonçalo Padrão**

Submitted for the degree of
Doctor in Biomedicine

Cognition and Brain
Plasticity Unit,
Department of Basic
Psychology - University
of Barcelona

Supervisor
**Antoni Rodríguez Fornells**
**PhD**

‘*To err is human; to forgive, divine.*’

By Alexander Pope

## ACKNOWLEGMENTS

So many ways, so many people, so many teachings... So many thanks!

This Thesis is dedicated to everyone who all over these years was JUST THERE, for the good and the bad things, present in the form of physical matter but also in spirit, supporting me and sharing the best they have, and who in one way or another found a way to live in my heart without paying a rent.

First I want to thank to my supervisor, and good friend, Antoni Rodriguez (Toni), for giving me this *second chance* and consequent opportunity of exploring the countryside of Catalonia with a piano keyboard on my back; running from wild pigs or crazy kids throwing me stones for no reason. Thinking on adventures like those ones make me laugh quite a lot, but also feeling profoundly grateful for your lesson and for trust in me since the first day we met. For your inspirational insights and for your curious sense of humor, for hearing me not as a boss but as a friend too, merci! From deep in side of my heart, Thanks a lot!

Then there is a great and inspiring team which I had the great pleasure to share my weekly routines of work, called BRAINVITGE F.C. (a team of beautiful beings and really talented researchers and students). I have many, many, many (exponential function) thanks to Josep Marco (aka The Wizard) for being the master crack of the data analysis but, above all, for being one the most honest, tough workers and good heart-souls that I've ever met, just inspiring (miedo me das por seres tan bueno tio!! Respect! Gracias de verdad!!). To LLuis and Ruth for their beautiful smiles and for being always so positive, nice and close to everyone, sharing their knowledge and good vibes. To David (David Cucureeeeeeellll :p) for all the reasons and all the reasons. My Thesis today is a reality because there are people like David swirling around in a cosmic place like Brainvitge. He is the man who makes things happen, and regarding my work, he really made it happen (Nen! Eres el Pxxx amo... gracias por el trabajo, por la amistad y compañerismo. Bless U!). Then there is a small group of mates who were present in my day to day life for longer periods of time (boring mates! So boring that became good friends). With them I had the pleasure of having good scientific discussions, have received with open heart their kind help and their insights. But with these guys I also have good moments of fun, whom I can trust and rely as brother usually do: Ernest and Clem, BROTHERS for LIFE!! To the rest of the Brainvitge crew my best respects (guys keep the vibe higher like this! You make all day to day frustrations of a PhD *career* so more easy... best atmosphere ever): Pablo (tio! que honor sentarme a tu lado, de veras. Y no sólo por tu santa sabiduría Matlábica pero por ser el verdadero Capitán del Brainvitge F.C. apoyando a cada uno

de nosotros y dando lo mejor de ti en cada momento). Diana, Viktória, Asia, Julià, Pau, Nuria, Helena, Estela, Claudia, Javiera, Jenny, Lucía, Pau, Ane, Ana Martínez, Marta, Júlia, Clara, Xavi, Dani, Ignasi and… Hope I'm not missing anyone. Familia, gracias!

To everyone who made part of my projects, giving their best: Mar, Borja, Laura, Aida.

To Joan Rodriguez for mastering the *Office* as anyone else on the planet Earth and for his important contribution on this Thesis.

To Virginia Penhune for being always so nice to me, for her comprehension, kindness and friendship. For her knowledge and teachings. Thanks a lot :)

To Pedro Soares (aka puto XibAx) for the cover of this book and for many suggestions regarding the graphic design.

To all friends that day to day take so good care of me, giving their love, showing their kindness and for teaching me so much.. For their love and inspiration, blessings!

Last, but not least, to my family, the reason why I have all the reasons of the universe for feeling so blessed and for being here right now, feeling ALIVE and HAPPY! To them, who had always protect me, who had always fight for our unity, for trust in me and for giving me all freedom to choose my path in accordance with my hearts' will (mamy, bro y abuelos, Love YAAA!) all the love, peace and light of the universe! Gratidão, de coraÇão.

Once again,
blessings, smiles, and true love
BOOM!

Gonçalo (piskalhao)

## SUMMARY

To err is certainly human. Detect and correct our errors is fundamental during our interaction with the outside world. Therefore, understanding the nature of the brain mechanisms involved in the flexible evaluation of human action and the adaptive changes that follow behavioral imperceptions is a basic goal of modern cognitive neuroscience.

The study of the brain mechanisms of error-monitoring has advanced enormously during the last two decades, mostly due to the discovery of specific electrophysiological signals and neural networks that are sensitive to error commission, but also to conflicting, unexpected and undesired events, all requiring the implementation of cognitive control processes in order to optimize performance. Neuroimaging studies, for instance, have associated error-monitoring with the activity of a widespread network of brain regions, wherein the medial prefrontal cortex is a key neural hub for regulative aspects of action monitoring and cognitive control. Electrophysiological studies have also identified a family of negative ERP signals in medial-frontal regions which appear to be mainly orchestrated by neural oscillatory theta activity. This field has provided the grounding for a very interesting research program regarding high-order cognitive control, decision-making and learning processes.

It is worth mentioning, though, that most of this research program has been mainly focus on the examination of action slips in fairly simple force-choice reaction time paradigms. In these contexts errors reflect no deliberated actions caused by perceptual or attentional lapses. However, in real life situations error forms are so widespread and its causes so diverse that a crucial challenge for cognitive neuroscientists concerns the development of methods and paradigms that allow the study of the neural bases of error-monitoring in broad ecological contexts that reproduce the complexity of everyday life situations in which humans are likely to commit errors.

The present dissertation aims at providing new alternatives and contributions regarding this issue by addressing novel questions, developing new toolkits and bringing new ideas to study well described neural dynamics of error-monitoring in more extended and ecological contexts in which humans interact. Throughout this research I have combined electrophysiological tools, fundamentally event-related potentials (ERPs) and time-frequency analysis, with novel experimental paradigms in order to provide answers to questions that all over these years have remained elusive and difficult to address experimentally.

This Thesis is composed by four studies which taking together put forward for consideration several ideas.

First, there is a substantial amount of visual inputs that are processed outside the focus of overt attention, and not available for conscious access, that still activate mechanisms in medial prefrontal control networks related to conscious and attentional processes. Neural theta oscillatory activity may stand as a neurobiological mechanism by which the medial-frontal networks monitor and regulate inappropriate actions that are automatically triggered by environmental information to which we remain oblivious.

Second, practice leads to functional changes in neurophysiological signatures associated with error-monitoring and error-awareness processes, which are crucial during the acquisition of new motor skills and learning.

Third, self-generated errors and errors related to agency violations are evaluated by distinct neural networks. The medial-frontal cortex is crucial for the evaluation of the correctness of ones actions while the parietal cortex seems to be more involved in providing a coherent sense of the agency, or sense of control, over ones actions.

Finally, different thresholds of error-tolerance in humans are related to different decisional processes and distinct patterns of cortical activity during the monitoring of redundant error feedback information in contexts involving rule-based decisions. These differences may reflect the externalization of distinctive cognitive schemas and standards of self-reinforcement to cope with errorful information in contexts requiring complex decision-making processes.

I believe that the findings forward in this dissertation are important to validate current neurophysiological evidences and theories regarding human error processing and cognitive control processes and may offer a great contribution to understand the extent and depth to which the human error-monitoring system can be studied extended and ecological contexts.

# RESUMEN

Errar es humano. Detectar y corregir nuestros errores es fundamental en nuestra interacción con el mundo exterior. Por lo tanto, la comprensión de los mecanismos cerebrales implicados en la evaluación de nuestras acciones y de los cambios adaptativos que siguen imperfecciones en nuestra conducta es un objetivo básico de la neurociencia cognitiva moderna.

El estudio de los mecanismos cerebrales implicados en el procesamiento de los errores ha avanzado enormemente en las últimas dos décadas debido al descubrimiento de señales electrofisiológicas específicas y redes neuronales sensibles a la comisión de errores, pero también a eventos conflictivos, inesperados y no deseados, que requieren la implementación de conductas compensatorias y procesos de control cognitivo con el fin de optimizar el rendimiento.

Estudios de neuroimagen, por ejemplo, han descubierto que la detección de errores está asociada a la actividad de una amplia red de regiones cerebrales, en que se destaca la corteza medial prefrontal. Estudios electrofisiológicos, por otra parte, han identificado una familia de potenciales evocados (PE) negativos en regiones medio-frontales que son coordenados por actividad neuronal oscilatoria de baja frecuencia en theta. Estos hallazgos han desvelado varias dinámicas relacionadas con los procesos de control cognitivo, toma de decisiones y aprendizaje orquestados por nuestro cerebro.

Merece la pena mencionar, sin embargo, que la mayor parte de esta investigación se ha centrado en la exploración de errores en tareas de tiempo de reacción que requieren acciones bastante simples. En estos contextos los errores habitualmente reflejan acciones no deliberadas, descuidos (*'action slips'*) resultantes de lapsos atencionales o perceptivos. Pero en la vida real los errores tienen varias formas y sus causas pueden ser tan diversas que es vital el desarrollo de paradigmas más ecológicos que reproduzcan la complejidad de la vida cotidiana en la que los seres humanos son propensos a cometer errores.

La presente tesis doctoral tiene como objetivo proporcionar nuevas alternativas y contribuciones con respecto a este tema, abordando nuevas cuestiones y desarrollando nuevos paradigmas para estudiar la compleja dinámica de los mecanismos cerebrales relacionados con la monitorización de errores y la consecuente implementación de acciones adaptativas en contextos más ecológicos y generales que abarquen la complejidad de nuestras intrincadas (yo diría complejas) interacciones con el ambiente. A lo largo de esta investigación he combinado herramientas clásicas en neurociencia cognitiva de electrofisiología, PEs y análisis tiempo-frecuencia, con nuevos paradigmas experimentales con el fin de dar respuestas a algunas preguntas que en todos estos años han permanecido esquivos y difíciles de abordar experimentalmente.

Esta tesis está compuesta por cuatro estudios que en conjunto aportaran nuevas ideas hasta el momento intangibles.

En primer lugar, se demuestra que hay una cantidad sustancial de información procesada fuera de nuestro foco atencional y no disponible a nuestra consciencia, que influencia nuestras acciones y, consecuentemente, puede activar mecanismos de la corteza medial prefrontal tradicionalmente relacionados con procesos conscientes y atencionales de control. Se propone, además, que las oscilaciones neuronales en theta constituyen un mecanismo neurobiológico adaptativo por el cual las redes medio-frontales de control cognitivo supervisan y regulan las acciones inapropiadas que se activan de forma automática por una gran cantidad de información ambiental no atendida pero aun así relevante.

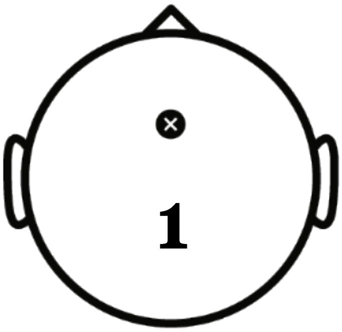
En segundo lugar, se explica como la práctica y el entrenamiento de una nueva habilidad conduce a cambios funcionales en los mecanismos neurofisiológicos asociados a procesos de monitorización automática de errores y procesos inherentes a la evaluación consciente de errores, funciones cerebrales cruciales en la adquisición de nuevas habilidades motoras y del aprendizaje.

En tercer lugar, errores autogenerados y errores relacionados con disfunciones en el control de agencia de nuestras acciones son evaluadas por redes neuronales distintas. La corteza medial-frontal es crucial en la evaluación de nuestras acciones mientras que la corteza parietal parece estar más involucrada en analizar la resultante experiencia de agencia y nuestra experiencia de control sobre las acciones que deliberadamente efectuamos.

Por último, diferentes umbrales de tolerancia al error en los seres humanos están relacionados con diferentes procesos de toma de decisiones y distintos patrones de actividad cortical relacionados con la evaluación de *resultados* negativos informativos sobre el resultado de nuestro desempeño. Estas diferencias pueden expresar la externalización de distintos esquemas cognitivos y normas de auto-refuerzo que los individuos asumen para hacer frente a información conflictiva o redundante en contextos propensos a errores que requieren procesos de toma de decisiones complejas.

Las ideas presentadas en esta tesis son importantes para validar pruebas neurofisiológicas actuales y teorías sobre el procesamiento de errores y adyacentes procesos de control cognitivo. Además ofrecen una gran contribución para futura investigación en la medida en que permiten comprender el alcance y la profundidad con la cual los sistemas cerebrales implicados en la monitorización de nuestras acciones pueden estudiarse de una forma más ecológica.

# Contents



# 1

# Introduction

## 1.1 Errors in human action

Errors are ubiquitous in human action. Although we try hard to avoid them during the accomplishment of different tasks during our everyday life errors still occur and their costs can range from mild annoyance (e.g. pour salt in the coffee cup) to devastating consequences. Therefore our ability to monitor our actions and detect errors may constitute one of the most important aspects involving adaptive and intelligent behavior and learning.

Errors can occur at many stages of performance, be very different in nature and, consequently, imply different regulatory (or corrective) actions to avoid them when possible. For instance, errors are quite common during early stages of skill acquisition, e.g., a pianist training a new musical piece from J.S. Bach, and may contribute to improvements in performance. Notwithstanding, errors still prevail during expert performance. In this last case, however, for both the observer and the performer it can be very surprising that someone considered under many criteria as an expert makes an unexplainable or an unbelievable error. Yet, events like famous and talented football players missing a penalty shoot in a crucial game have remained in the memory of some of us for long time, for instance when the Italian player Roberto Baggio missed a penalty shoot crowning Brazil world champion. At other level of the iceberg (probably at the very top, regarding the terrible costs that human errors can have), remote incidents like the disaster at Los Rodeos Airport on Tenerife, Spain, on 1977 have also been, unfortunately, over and over described as the consequence of implausible erroneous decisions taken by human expert operators (Casey, 2006; Wiegmann & Shappell, 2003; Woods, 1994). Therefore a better understanding of the different factors contributing to human failures in these contexts might not only mitigate its potential disastrous consequences but may also inform on the underpinnings of performance across domains in psychology and neuroscience (Reason, 1990).

The interest around the factors that contribute to human errors, the way they are interpreted and evaluated by the human mind, and their impact on intelligent behavior, has increasingly drawn the attention of cognitive and experimental psychologists and, more recently, of cognitive neuroscientists. Because errors reflect a salient marker that performance has broken down, the investigation of the cognitive and neural underpinnings of error processing have for long time being assumed by many theories of executive or cognitive control (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Holroyd, Nieuwenhuis, Mars, & Coles, 2004; Miller & Cohen, 2001; Norman & Shallice, 1986; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004). Solving difficult, novel, or complex tasks, overcoming error (or error-prone) actions and correcting them, all require a high degree of cognitive control. In this sense, it has long been recognized that errors provide vital information for regulative and compensatory control process in order to optimize goal-directed behavior.

In the mid-1960s Rabbit and colleagues first demonstrated that errors are often followed by compensatory behaviors including fast error-correction processes or slowing down the onset of actions subsequent to erroneous actions (Rabbitt, 1966a; Rabbitt, 1966b). Since the work of Rabbit and colleagues the study of the brain mechanisms underlying error-monitoring advanced enormously. Thanks to the development of electrophysiological and neuroimaging methods, human and monkey neurophysiology and neuroanatomical and neuropsychological analysis, the field has burgeoned and provided the grounding for a very interesting research program regarding high-order executive control mechanisms. At the present moment error-monitoring and control processes in a wide variety of contexts are believed to be accomplished by a generic and flexible system that is mainly localized in areas of the medial-frontal cortex (MFC), and specifically in the more posterior part encompassing the anterior cingulate cortex (ACC), the pre supplementary motor area (pre-SMA) as well as adjacent SMA. Hereafter, I will refer to these brain structures involved in action (or performance) monitoring and adaptive control of behavior as the posterior medial-frontal cortex, pMFC (*see* Fig. 1)[1].

In the following sections I will review some of the most important empirical data acquired during the last years of research in cognitive neuroscience regarding error-monitoring and related cognitive control processes. First, a brief description of the most common experimental paradigms used to study human errors in the laboratory setting will be presented (section 1.2). Then I will review a set of behavioral (1.3) and neuroimaging studies (1.4), with a main focus on electrophysiological findings, which taken together implicate the pMFC during strategic aspects of cognitive control of motor behavior, being especially important in the evaluation of errors, response conflict, unexpected and salient events, and consequent remedial processes. In the section 1.5, important theoretical accounts of error-monitoring and cognitive control are briefly exposed: (i) the Comparator model, (ii) the Conflict-monitoring model, and (iii) the Reinforcement Learning theory of error processing. Finally, I will introduce four different lines of research that have been developed somehow independently of the cognitive neuroscience approach of error processing, but which I considered of crucial importance to understand the extend and depth to which error-monitoring processes can be studied in more ecologic contexts in which humans interact. This very last section (1.6) introduces the four studies that constitute the core of the present dissertation thesis.

---

[1] This nomenclature has been adopted in previous research, see for example a review paper from Ullsperger et al. (2014)

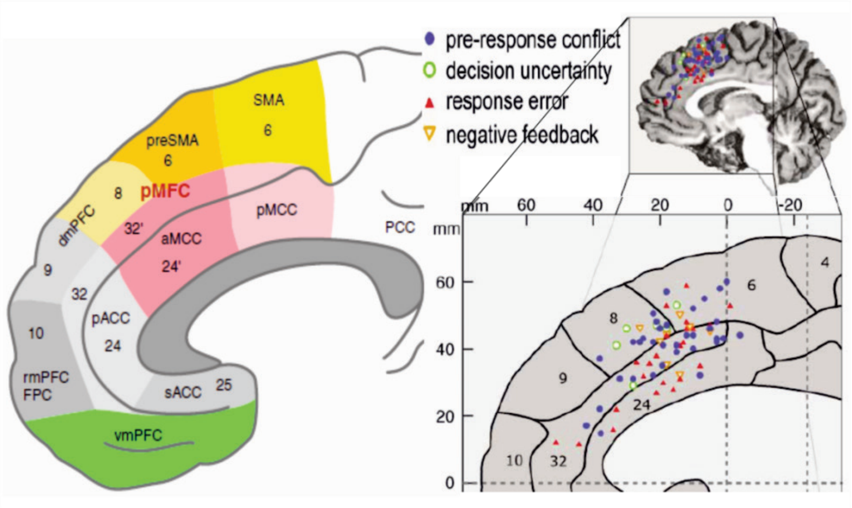


**Figure 1:** On the left side is illustrated the view of the sub-regions of the medial-frontal cortex. The posterior medial frontal cortex (pMFC) is a nomenclature that recently has being used to describe a broader region of the medial-frontal cortex including neurons of the pre-supplementary motor area (pre-SMA, 6), the anterior and medial cingulate cortex (aMCC, 24 and 32), and adjacent to these areas the suplementary motor area (6); together these regions comprise the neural locus of the action monitoring system. vmPFC, ventromedial prefrontal cortex; rmPFC, rostromedial prefrontal cortex; dmPFC, dorsomedial prefrontal cortex; sACC and pACC, respectively, subcallosal and pregenual anterior cingulate cortices; pMCC, posterior midcingulate cortices (adapted from Ullsperger et al., 2014). On the right side of the figure is represented a meta-analysis of fMRI studies showing activations in regions comprising the pMFC during error commission, pre-response conflict, uncertainty and feedback processing (adapted from Ridderinkhoff et al., 2004).

## 1.2 Experimental paradigms to study error-monitoring and cognitive control in laboratory settings

Paradigms to study error-monitoring and cognitive control processes come in many flavors. However, a common feature across them is that errors are usually induced by experimental conditions associated with increased conflict or interference during the response generation process. Often this conflict is induced by the competition between events that stimulate prepotent, but inappropriate, response tendencies (the conflicting stimuli) and events that generally specify the correct response to be executed according to the task goal.

One of the most widely used paradigms is the Erikson Flanker task (Eriksen & Eriksen, 1974), wherein participants are required to make fast discriminative responses to a target stimulus that is flanked by other distractor stimuli which should remain unattended. Distractors can be either congruent, i.e. associated with the same response as the target (e.g. <<<<<), or incongruent, conflicting with expected correct response (e.g. <<><<) (Fig. 2). Another task frequently used is the 'Stroop Task'

(Stroop, 1935). In this task subjects are presented with the name of a color printed in colored ink (Fig. 2) and must identify the color of the ink as fast and accurately as possible. In incongruent (or high conflicting) conditions, the color's name differs from the ink color (e.g. the word "RED" printed in green ink) whereas in congruent (low conflict) conditions the color name matches the ink color. The logic beneath these tasks is that incongruent trials typically pre-activate automatically conflicting action plans, since processing of the incongruent, task-irrelevant, events activates the alternative (usually the incorrect) response to the target stimuli. Therefore, to avoid errors participants are required to override task-inappropriate response tendencies triggered by the conflicting condition. At the behavioral level, higher conflicting conditions (e.g. when the distractor conflicts with the correct response pattern) are often accompanied by higher error rates and slower reaction times (Fig. 2) (Botvinick et al., 2001; Gratton, Coles, Sirebaag, Eriksen, & Donchin, 1988). These classical paradigms to a certain extent show the influence that unattended, but task-irrelevant, inputs may have over goal-directed behavior in tasks involving focused attention (Miller, 1991).

Other variants of these tasks, such as the Go/No-go or Stop-Signal tasks have also been used. In these tasks participants perform speeded responses to go-signals, but are instructed to refrain from responding when the go-signal is followed by a Stop or a No-go signals (Logan, 1994). Because the No-go or the Stop signal appears in a small proportion of trials, participants often show problems to inhibit the pre-programmed action planned triggered by the go signal incurring in higher error rates.

More recently other studies have cleverly used other type of tasks involving, for example, perceptual discrimination (Pavone, Marzi, & Girelli, 2009; Steinhauser & Yeung, 2010) or motor tracking tasks (Krigolson & Holroyd, 2006). In all of these tasks errors generally reflect action slips, or in other words, they result from actions that are carried out "unintentionally" (Norman, 1981; Norman, 1988), mostly due to perceptual lapses, attentional fluctuations, loss of activation or maladjustments in control and executive behavior (Botvinick et al., 2001; Eichele et al., 2008; O'Connell et al., 2009). However, as it will be discussed in a later section of this manuscript (see section 1.6), out of the laboratory setting errors can be very different in nature and not necessarily related to attentional or control fluctuations in the nervous system. Therefore, it is important to invent new experimental paradigms in order to investigate performance errors in broad ecologically valid contexts in which humans interact in their everyday life.

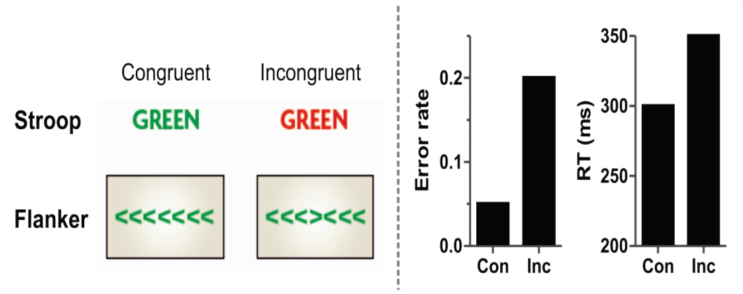


**Figure 2:** On the left side, examples of tasks that involve response conflict. In the Stroop task (top), the subject must name the ink of the color, which can be written matching the ink color (Congruent trials) or be different from it (Incongruent trials). In the Flanker test (bottom), subjects make fast responses to the direction of the central stimulus which is flanked by stimuli with the same (Congruent trials) or opposing (Incongruent trials) directions. On the right side, Error rate and Reaction Time (RTs) as function of the congruency. Note the increase in error rate and RTs during the conditions of increased conflict, i.e. in the incongruent trials (simulated data).

## 1.3 Behavioral evidences of error-monitoring and compensation processes in humans

Seminal work by Rabbitt in the mid-1960s first raised the importance of a system responsible for error detection and compensation (Rabbitt, 1966a; 1966b). His results were remarkable by showing that performance errors, which were often accompanied by emotional reactions of frustration, are usually followed by automatic compensatory (or regulatory) behaviors. During his observations in reaction time tasks, he showed that adaptive behaviors such as slowing down the production of correct responses in subsequent trials following an error, a process called post-error slowing (PES) or increase the accuracy in trials following an error were very common (Rabbit, 1966a; 1966b; *see* also Laming, 1979). Recently, it has been demonstrated that PES may be related with the decrease of activity in motor areas after errors are committed (Danielmeier & Ullsperger, 2011; King, Korb, von Cromon, & Ullsperger, 2010). The decrease in motor activity following errors presumably reflects adjustments of the motor threshold, or motor inhibition that are triggered in order to regulate the urgency or caution with which a given task is performed. The PES effects has been observed in the most classical experimental tasks, for instance in the Flanker (Cavanagh, Cohen, & Allen, 2009; Debener et al., 2005; Marco-Pallares, Camara, Munte, & Rodriguez-Fornells, 2008), the Stroop (Gehring & Fencisik, 2001), or the Simon tasks (Danielmeier, Eichele, Forstmann, Tittgemeyer, & Ullsperger, 2011; King et al., 2010); but also in tasks involving speech production (Levelt, 1989), motor learning (Angel & Higgins, 1969), skilled music performance (Ruiz, Jabusch, & Altenmuller, 2009) or skilled typewriting (Logan & Crump, 2010). Recent evidences suggest that cautious behaviors following adverse events or inappropriate action sets, such as PES, represent a common behavioral feature among mammals, being observed even in rodents (Narayanan, Cavanagh, Frank, & Laubach, 2013).

Furthermore, we humans do not only 'hold our horses' after making errors but we also correct them quite naturally in an automatic fashion (Rabbit, 2002). Error correction is indubitably one of the fastest cognitive processes; subjects tend to correct their errors even when they are instructed to suppress such responses (Rabbitt & Rodgers, 1977). Cooke and Diggles (1984), for instance, by recording electromyographic activity *EMG*[2] while participants performed a visually guided tracking task involving arm movements showed that the onset of corrective muscle movements started as early as 20-40 ms after the initiation of the erroneous movement, suggesting that error-correction mechanisms are probably independent from proprioceptive external feedback of the motor output. It is possible that error-monitoring and correction may develop in parallel and are probably dissociated from the conscious experience of error detection. In contrast with the velocity in which errors are corrected, conscious error detection and error signaling (when for example subjects are required to deliberately acknowledge their errors by signaling their errors with key presses) appear to be much slower and effortful, ranging from 500 to 800 ms after error commision (Rabbitt, 2002).

Interestingly, Coles and colleagues (Coles, Gratton, Bashore, Ericksen, & Donchin, 1985) revealed that on a proportion of correct trials there are subthreshold muscle twitches (recording through the EMG) before the response is given that are often observed in error responses too (see also Burle, Roger, Allain, Vidal, & Hasbroucq, 2008). This finding suggests that few correct responses are accompanied by partial error activity that is evaluated on-line and aborted before reaching the overt response threshold. PES effects have been also shown after partial errors, although this effect is much smaller than for overt error responses(Allain, Burle, Hasbroucq, & Vidal, 2009). These observations suggest the existence of an on-line control mechanism that allows the nervous system to detect, stop, correct error response tendencies, and implement behavioral adjustments. However, the question as to whether those covert incorrect muscle activations (error-monitoring processes) are consciously perceived remains as an important open question.

Another common finding in laboratory tasks involving cognitive interference is that in conflicting trials error rates are typically higher and subjects also tend to reduce the speed of their correct responses. To a certain extent the RT slowing down in trials of increased conflict can be also understood as an on-line *reactive control process* which is ensued to overcome erroneous response tendencies. The behavioral effects resulting from conflicting events are not just limited to the current trial

---

[2] EMG is a technique for recording the electrical activity produced by skeletal muscles, which detects the electrical potential changes generated by the muscle cells.

(reactive control processes), but they also affect performance in future behavior, in which they are manifested as a behavioral improvement if the subject is faced with previous conflicting events again (also referred to as proactive control processes) (Braver, 2012)[3]. For instance, RTs in high-conflict trials that are immediately preceded by another high-conflict trial are shorter than those in high-conflict trials that are immediately preceded by a low-conflict trial (Botvinick et al., 2001; Gratton, Coles, & Donchin, 1992; Kerns et al., 2004). This facilitative effect of previously experienced conflict has been demonstrated in a range of different tasks and supports the idea that response conflict in a given situation triggers control adjustments to reduce conflict and optimize performance in a subsequent trials.

Together these findings point to the existence an error or performance monitoring system which aids the implementation of compensatory and adaptive actions in order to optimize performance. In the next section of this dissertation I will overview some of the neuronal mechanisms underlying such behavioral adaptions after errors or other performance problems, such as high response conflict, novelty, or task difficulty.

## 1.4 A neural system for error-monitoring and compensation

### 1.4.1 Electrophysiological evidences of error-monitoring in humans

The first insights into the neural basis of an error-monitoring system in the human brain derived from experiments using *event-related potentials (ERPs)*[4]. In the early 1990s, researchers observed, for the first time, that erroneous responses committed by human participants in reaction time tasks were associated with a fast negative ERP spatially distributed over frontocentral regions of the scalp (Fig. 3B). The component is today known as the error-related negativity (ERN) (Falkeinstein, Hohnstein, Hoormann, & Blanke, 1990; Gehring, Goss, Coles, Meyer, & Donchin, 1993). The ERN component is typically observed in response-locked waveforms of the on-going EEG as a difference between averaged error and correct responses and its onset occurs at or slightly after the moment of response initiation, peaking at about 50-100 ms after response onset (Fig. 3A). The precise latency of this component depends on the time-locking event. Most studies use button-press responses to signal

---

[3] reactive control process make reference to behavioral adaptations in the current trial while regulatory or proactive cognitive control processes make reference to preparatory control processes after the experience of conflict/errors (i.e. in the next trial).

[4] ERPs reflect changes in the EEG signal (measured in µV) that are evoked by occurrence of a discrete event. ERPs responses are usually obtained by the average in the time-domain of multiple trials containing the event of interest. ERPs is a technique particularly suited to investigating the time course of fast cognitive processes.

the onset of erroneous movements, others, however, record EMG activity to time the moment of "erroneous" muscle activations (Burle et al., 2008; Gehring et al., 1993).

It is been hypothesized that the ERN originate in the pMFC, specifically in regions encompassing the anterior cingulate cortex (ACC) and premotor supplementary areas (pre-SMA) (Debener et al., 2005; Dehaene, Posner, & Tucker, 1994; Ullsperger, Danielmeier, & Jocham, 2014; Yeung, Botvinick, & Cohen, 2004) (Fig. 3C). Very recently, a study using intracerebral recordings in humans undergoing surgical evaluation for epilepsy has revealed local field potentials in the supplementary motor area (SMA) approximately 150 ms after error-related EMG activations were recorded (Bonini et al., 2014).

The fast onset of the ERN have led some authors to suggest that this neural signal is dissociated from the subjective experience of error awareness (or error detection) (Endrass, Franke, & Kathmann, 2005; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001; Overbeek, Nieuwenhuis, & Ridderinkhof, 2005). In turn, conscious error detection has been associated with a centroparietal positive ERP which characteristically exhibits slow sustained activity (from 300 to 500 ms after response errors) following the ERN response (Murphy, Robertson, Allen, Hester, & O'Connell, 2012; Nieuwenhuis, et al., 2001; O'Connell et al., 2007; Steinhauser & Yeung, 2010). This ERP signal, called error-related positivity (Pe), in terms of latency and spatial distribution resembles the P3 signal as a neural response to salient target stimuli in oddball tasks (Ridderinkhoff et al., 2009; Abel & Donchin, 2009). Interestingly, the offset time of the centroparietal Pe (~ 500 ms after error onset) is consistent with the time required by young participants to signal (acknowledge) their errors, more or less 500-700 ms (Rabbit, 2002). Recently it has been suggested that the centroparietal Pe reflects a neural computation associated with the accumulation of internal evidences that an error as occurred (Steinhauser & Yeung, 2010) leading to the conscious experience of error commission.

Independently of the degree of error awareness, the ERN amplitude is sensitive to the value of the outcomes (Bernstein, Scheffers, & Coles, 1995) and to task instructions signaling participants for the importance of errors in a given context (Gehring et al., 1993). Moreover, the ERN amplitude has been related to compensatory processes, such as error correction (Gehring et al., 1993; Rodriguez-Fornells et al., 2002; Burle et al., 2008), error force (Gehring et al., 1993) and PES (Gehring et al., 1993; Debener et al., 2005).

From the literature in cognitive neuroscience, it is quite well established that the ERN echoes neural activity of a highly flexible system involved in the rapid evaluation of ongoing events, being particularly sensitive to inappropriate or conflicting action-sets and regulative aspects of cognitive and motor control (Botvinick et al., 2001; Holroyd & Coles, 2002; Holroyd et al., 2004). A wide variety of psychological tasks using visual, auditory or tactile stimuli have reported error-related ERN activity independent of the response effector (for a review *see* Gehring et al., 2012), indicating a higher flexibility of the system that generates the ERN.

Although some errors can be detected right after the erroneous action is initiated, there are many situations in which we have to wait for external information to detect whether a certain action was accomplished successfully. For example, when playing darts we can only evaluate how good our shoot was when the dart hits the board, informing the points we have gained. Yet, depending on the experience of the performer, the evaluation of the likelihood of a good shoot can be predicted based on the fluency with which the action of throwing the dart is initiated independently of external information (Aglioti, Cesari, Romani, & Urgesi, 2008; Ruiz et al., 2009).

Many ERP studies have observed that external feedback signaling about unexpected, negative outcomes and incorrect performance are followed by a frontocentral negative ERP component very similar in terms of the spatial distribution to the ERN (Gehring & Willoughby, 2002; Holroyd & Coles, 2002). This ERP, termed feedback-related negativity (FRN), occurs approximately 250 ms after feedback presentation (Fig. 3A, middle panel). The FRN it has been observed in time-estimation tasks (Miltner, Braun, & Coles, 1997), gambling tasks (Gehring & Willoughby, 2002; Marco-Pallares, 2007) and tasks of associative learning (Holroyd & Coles, 2002). Evidences from electrical source localization methods point that the source of the ERN and FRN overlap, partially, in the same neural structure of the brain, the pMFC (Gruendler, Ullsperger, & Huster, 2011; Muller, Moller, Rodriguez-Fornells, & Munte, 2005) (Fig. 3C). The FRN is also followed by a later positive deflection on the EEG signal, called the P3 component.

Both the Pe (following the response ERN) and the P3 (following the feedback FRN) share similar features such that recently it has been proposed that both signals may reflect similar neural computations associated with the updating of relevant (i.e. motivational salient) actions or external feedbacks to current task contexts involving the formation of internal decisions (Fisher & Ullsperguer, 2013;O'Connell, Dockree, & Kelly, 2012; Hillyard, Squires, Bauer, & Lindsay, 1971), as for example the accumulation of internal evidences that an error has occurred (Steinhauser & Yeung, 2010) and the generation of new actions may be required.

As the response ERN, the feedback FRN shows a graded sensitivity to the value of outcome (Gehring & Willoughby, 2002; Nieuwenhuis, Yeung, Holroyd, Schurger, & Cohen, 2004; Oliveira, McDonald, & Goodman, 2007); its amplitude is generally higher on more unexpected outcomes, indexing a *prediction error signal*[5], i.e. the difference between predicted and obtained action outcome (Chase et al., 2011;

---

[5] Prediction error signal: the difference between the outcome obtained when performing an action and the expected outcome of performing that action.

Holroyd & Coles, 2002; Mas-Herrero & Marco-Pallares, 2014). Moreover, the amplitude of the FRN seems to be influenced by the degree of participant's involvement with the task, which suggests that the evaluative process indexed by the FRN is sensitive to the motivational significance of ongoing events (Yeung et al., 2005).

Furthermore, error-related brain activity has also been verified when people observe error events made by other agents. Errors performed by other agents (i.e. observational errors) elicits a frontocentral negative ERP deflection similar the feedback-related ERN, which is localized in the pMFC (van Schie et al., 2004). These findings put forward that the neural mechanisms involved in the monitoring of one's own actions are also activated when monitoring other's actions, offering an explanatory view on how our brain detect other persons' mistakes and therefore may use this information to adjust our own behavior by observational or vicarious learning.

Another set of ERP findings in reaction time-tasks involving response conflict have suggest that conflicting, error-prone events are related to another frontocentral negative component which peaks 300 ms after the presentation of the conflicting stimuli and is followed by a later positive P3 deflection (Folstein & Van Petten, 2008; Yeung et al., 2004; van Veen & Carter, 2002; ). This negative component, called N2, has a topographical distribution similar to the ERN and the FRN and is originated is regions of the pMFC, more specifically in the ACC (van Veen & Carter, 2002; Yeung et al., 2004) (Fig. 3, right panel). The resemblance between the ERN and N2 components led some authors to claim that both cortical responses reflect similar functional mechanisms involved in conflict monitoring and cognitive control: the N2 involved with the monitoring of conflict between competing response tendencies prior to the execution of correct responses; and the ERN indexing post-response conflict processes following erroneous responses that could not be overridden successfully (Botvinick et al., 2004; van Veen & Carter, 2002; Yeung et al., 2004). The N2 has also been found during inhibition of preactivated motor responses in Go/NoGo and Stop signal tasks (Donkers & van Boxtel, 2004; Folstein & van Petten, 2008). The N2 signal in successfully inhibited NoGo events can be also partially explained as a mechanism of conflict-monitoring triggered by the No-go cue that automatically indicates the need of inhibit the inappropriate response tendency.

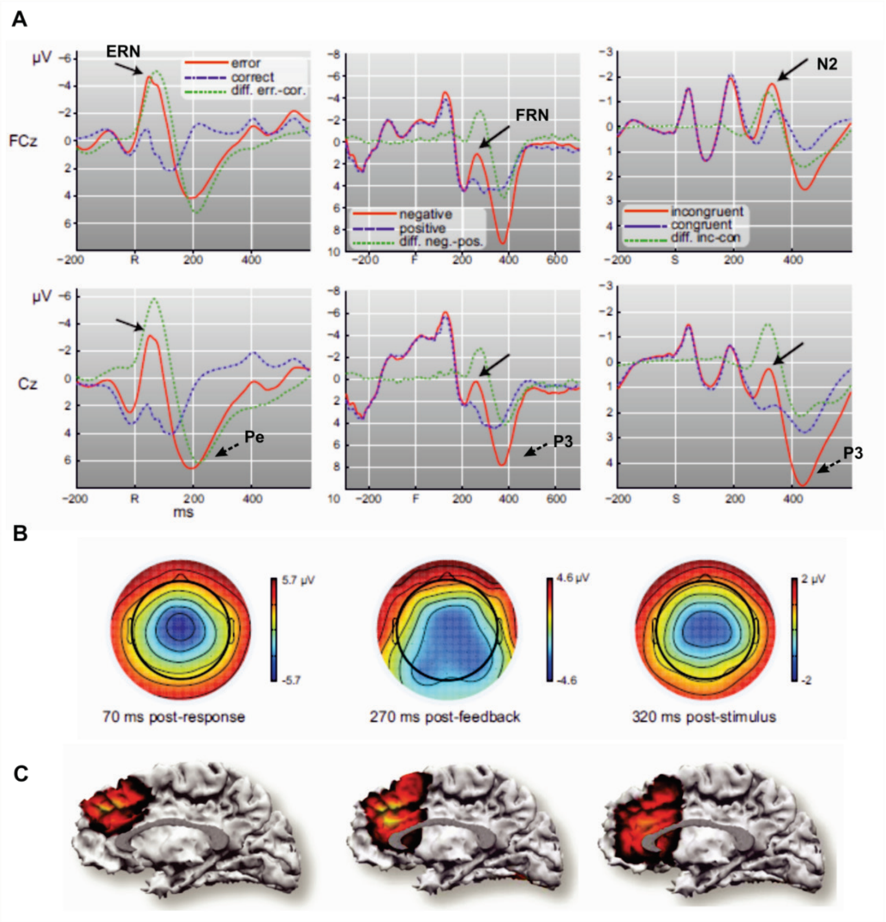


**Figure 3: (A)** Event-related potentials (ERPs) associated with error commission (ERN, response-locked), negative feedback (FRN, feedback-locked) and pre-response conflict (N2, stimulus-locked) at midline electrodes FCz and Cz (medial-frontal region). Also note the positive ERP defections following the ERN (the Pe component), and the FRN and N2 (the P3 component). These positive ERP deflections are thought to potentially manifest different aspects of performance monitoring, such as the awareness of relevant actions and events. **(B)** Topographical maps during the moment of peak activity of the ERN (70 ms), FRN (270 ms), and N2 (320 ms). **(C)** Source localizations of the ERN, FRN, and N2 showing the activation of the MFC(adapted from Ullsperger et al., 2014).

To sum up, evidences from ERP studies demonstrate that frontocentral negative ERPs such as the ERN, FRN and N2 may reflect the output of the pMFC during the monitoring of errors, negative outcomes, or error-prone conflicting situations that need to be compensated. On the other hand, slow positive ERP signals over centroparietal regions of the scalp, such as the Pe and the P3, seem to reflect neural computations associated with the updating of relevant events - response errors, novel or unexpected feedbacks, or conflicting information - and the consequent awareness of those events to current task contexts leading, perhaps, to deliberate changes in on-going behavior.

All these neural signals to certain extent reflect different aspects of performance monitoring, though they seem to clearly tap critical aspects of action monitoring and cognitive control in response to a broader context of plans and goals.

### 1.4.2 The pMFC: a neural hub for action monitoring and cognitive control

In the last years we have assisted to an explosion of research on the role of the pMFC, and specifically the ACC, in cognition which has led to a proliferation of theories about its function. Most of these theories widely agree that this region is involved in many aspects of action monitoring, such as error and conflict monitoring, feedback evaluation and uncertain decisions (Ridderinkhof et al., 2004), and a key neural interface for action selection and cognitive control (Botvinick, Cohen, & Carter, 2004; Holroyd & Coles, 2002; Holroyd et al., 2004).

The ACC, which lies on the medial surface of the frontal lobes, receives many inputs from the limbic system and from cortical structures including the prefrontal cortex and motor cortex and, in turn, project many neurons to motor brain regions including the basal ganglia, the primary and supplementary motor areas, and the spinal cord (Paus, 2001). The convergence of fronto-cortical and cortico-subcortical neural connections in the ACC, distinguishes this region as a powerful hub to regulate the interaction between cognition and motor control according to different motivational states of the organism (Paus, 2001; Shackman et al., 2011); a functional node to translate internal intentions to actions and, as some authors have suggested, a motor control filter that evaluates the “appropriateness of ongoing events” during goal-directed behavior (Holroyd & Coles, 2002).

The involvement of the pMFC in error monitoring was probably first demonstrated in a study by Niki and Watanabe (1979) in which changes in the activity of single ACC neurons were recorded when a monkey made errors. Years later, error-related potentials were observed in the anterior cingulate motor region during monkeys’ errors in simple response tasks (Gemba, Sasaki, & Brooks, 1986; Gemba et al., 1986). Invasive recording techniques of cellular activity in animals have been fundamental to demonstrate the responsiveness of pMFC neurons to response errors and unexpected omissions of rewards (Ito, Stuphorn, Brown, & Schall, 2003; Shima & Tanji, 1998). In the study of Ito and colleagues (2003), for example, single-unit activity in the ACC was recorded while monkeys performed a saccade-countermanding task with the aim of dissociate neural signals responsive to erroneous actions, unexpected reinforcements, and response conflict. It was found that some neurons on the ACC were activated during errors (Fig. 4A) and, interestingly, half of which were also responsive to the omission of a earned reinforcement (Fig.4b). These neurons were not modulated by correct cancelled stop signal trials neither by correct trials with no-stop signals (Fig. 4C), suggesting that ACC neurons are mainly responsive in monitoring the consequences of actions.

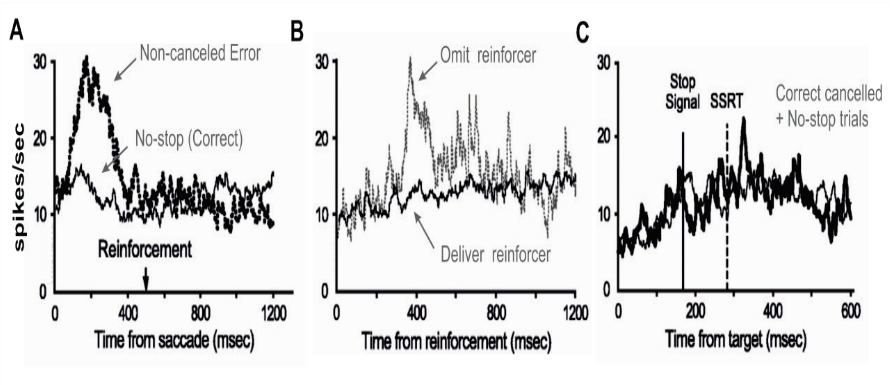


**Figure 4.** Single-unit activity recorded in the ACC while monkeys performed a saccade-countermanding task. **(A)** Neurons exhibiting phasic error-related activity: Activity for successful n-stop signal trials (thin solid line) and erroneous noncanceled trials (thick dotted line). Arrows indicate the time of reinforcement on no-stop signal trials. **(B)** error-related neurons responsive to the omission of a reinforcement. Activity when the reinforcement was delivered (solid line) or withheld (dotted line). **(C)** Activity aligned on target presentation for canceled trials (thick solid line) and no-stop signal trials (thin solid line). No differences between conditions show that these neurons do not signal conflict (adapted from Ito et al., 2003).

Several functional neuroimaging studies *(fMRI)* have also shown quite consistently the activation of a wide brain network comprising the pMFC (including the ACC and pre-SMA), inferior frontal gyrus (IFJ), dorsolateral prefrontal cortex (DLPFC), and the insular cortex, when participants make error responses in speeded choice reaction time tasks (Carter et al., 1998; Debener et al., 2005; Kerns et al., 2004; Marco-Pallarés et al., 2008; Ullsperger & von Cramon, 2001) (Fig. 5A).

In a seminal study, Debener and colleagues (2005) recorded simultaneously EEG and fMRI to observe trial-by-trial fluctuations of error-related activity in the human brain. Single-trial ERN modulations after errors were confirmed (Fig. 5B, left: note that the ERN was seeded on the rostral part of the cingulate cortex). Critical, their analysis of simultaneous EEG/fMRI measurements revealed that error-related FMRI activations correlated positively with single trial amplitudes of the ERN in rostral part of the cinculate cortex (pMFC), such that greater single-trial ERN amplitudes were associated with stronger BOLD responses in the pMFC (Fig. 5B, right). The authors suggested that the ERN amplitude reflects the putative activation of pMFC neurons involved in action monitoring and subsequent behavioral adjustments.

Several other neuroimaging studies have demonstrated crucial participation of the pMFC during situations involving pre-response conflict, negative feedback evaluation or decision uncertainty (Behrens, Woolrich, Walton, & Rushworth, 2007; Camara, Rodriguez-Fornells, & Munte, 2009; Carter et al., 1998; Kerns et al., 2004; Knutson, Adams, Fong, & Hommer, 2001; Paulus & Frank, 2006; Ridderinkhof et al., 2004; Yacubian et al., 2006).

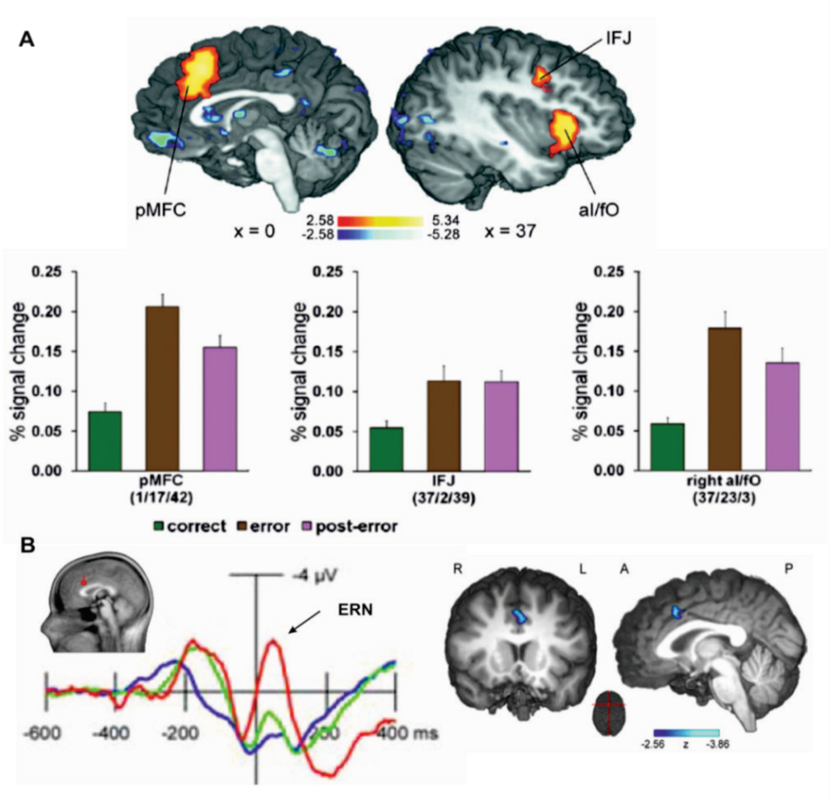


**Figure 5. (A)** fMRI Activation patterns in brain regions associated with error-monitoring. On the top, sagittal views show greater activity in error trials relative to correct trials (warm colors) and the reverse activation pattern (cool colors). Note the increase activity in the pMFC, Inferior Frontal gyrus (IFG) and Insular cortex (aI/oF). On the bottom is also indicated the BOLD percentage signal change on correct, error, and post-error trials in these regions (adapted from King et al., 2010). **(B)** Concurrent EEG and fMRI analysis identifying a network responsible for error-monitoring. *On the left*, the grand-average ERPs displaying the ERN (single-trial ERN extracted by independent component analysis) increase in error trials (red). Electrical source localization of the single-trial ERN identified a single dipole seeded in pMFC areas. On the right, correlations analysis showing that error-related fMRI signals correlated with single-trial amplitudes of the ERN in the pMFC, suggesting that the ERN reflects the putative activation of pMFC neurons during error-monitoring (adapted from Debener et al., 2005).

Moreover, error-related activity in the pMFC also predicts quite consistently PES effects (Kerns et al., 2004; Garavan, Ross, Murphy, Roche, & Stein, 2002; Marco-Pallarés et al., 2008; Danielmeier et al, 2011). In fact, lesion data has been important to show that regulatory processes depend on the integrity of prefrontal brain structures (Narayanan et al., 2013; Sheth et al., 2012). In a multimodal study by Sheth and colleagues (2012) using fMRI, human single-neuron recordings and integrating behavioral observations before and after human subjects underwent surgical cingulotomy, it was shown that conflict adaption, i.e. reduction of conflict costs in conflicting events preceded by higher conflicting events, was impaired after removal of the dorsal ACC.

In accordance with major theories of cognitive control, the pMFC following conflicting events, errors, or negative outcomes alerts the system to increase control and guide behavioural adjustments - possibly by increasing attention to the task or informing preparatory motor areas of the brain to correct, to adjust or to switch the output of the on-going actions (Botvinick et al., 2001; Holroyd & Coles, 2002; Kerns et al., 2004; Ridderinkhoff et al., 2004). However, top down control mechanisms may be mainly orchestrated by other brain structures such as the lateral (LPFC), or dorsolateral (DLPFC) prefrontal cortex (Botvinick et al., 2004; Cavanagh et al., 2009; Cohen, Botvinick, & Carter, 2000; Kerns et al., 2004). In 2000, McDonnald et al. first demonstrated dissociation between the ACC, responsible for evaluative or monitoring processes, and the DLPFC, responsible for maintaining top-down control during task performance. Using a modified version of the Stroop paradigm it was observed that the DLPFC, but not the ACC, was mainly active in response to instructions requiring greater top-down control, which suggests "the role of the DLPFC in representing and maintaining task demands needed for such control". In turn, the ACC was mainly activated during the response process on more conflicting trials replicating previous findings (Botvinick & Cohen, 1998; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Carter et al., 1998). Therefore it was suggested that while the ACC is responsible for monitoring the occurrence of errors or the presence of conflict, being crucial in providing feedback to the system as to when strategic processes must be more strongly engaged to adapt ongoing behavior, the DLPFC is responsible for strategic control adjustments, representing and maintaining the attentional demands of the task. A follow up study from the same group (Kerns et al., 2004) further showed ACC conflict and error-related activity predicts greater LPFC activity and adjustments in behavior (Fig. 6A).

In line with this ideas, it has also been shown that signals from the pMFC after errors can exert control over brain regions that are responsible for the perceptual processing of task-relevant stimulus features in the visual cortex (Danielmeier et al.,

2011) and brain regions implicated in motor execution (Danielmeier et al., 2011; King et al., 2010) in order to promote post-error adjustments (Fig. 6B).

Together these findings suggest that the pMFC is an important locus for action monitoring and especially sensitive for the computation of performance errors. Moreover, this region is not only involved in evaluative processes of action but mediates dynamic performance adjustments.

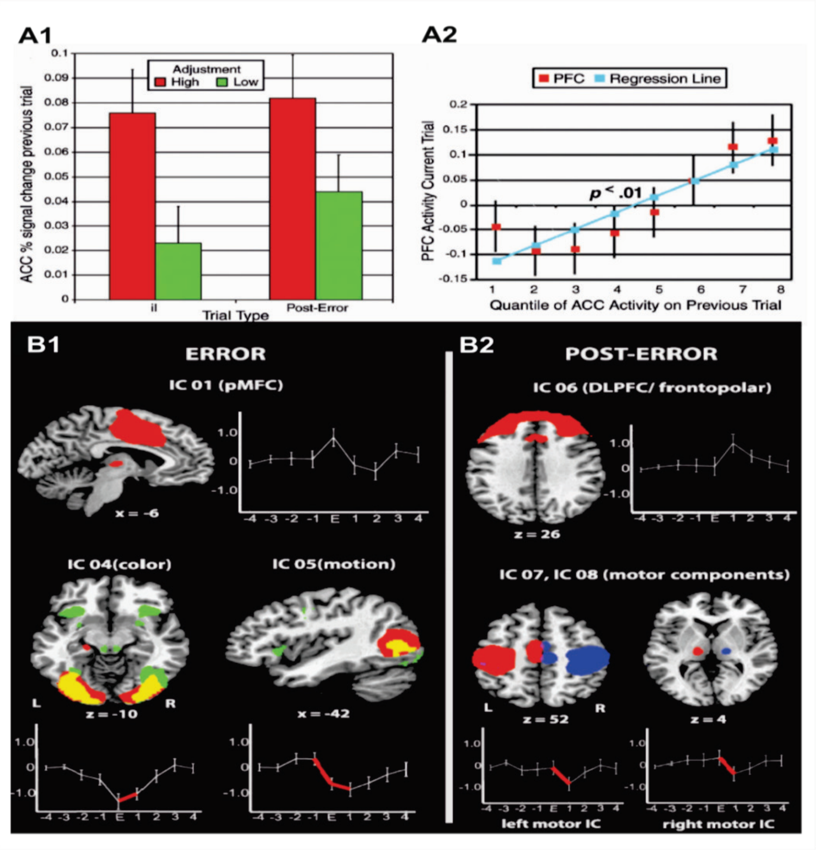


**Figure 6.** (**A**) ACC conflict and error-related activity predicts greater LPFC activity and adjustments in behavior. **A1.** ACC activity on the previous incongruent and error trials predicted behavioral adjustments on the next trial. High adjustment means fast RT on iI trials and slow RT on post-error trials (i.e. post-error slowing). Low adjustments are reflected in slow RT on iI trials and fast RT on post-error trials. **A2.** ACC activity on previous conflict (incongruent) and error trials predicts PFC (lateral) activity on the current trial (adapted from Kerns et al., 2004). **(B)** Error-related pMFC activity predicts subsequent changes in control, perceptual and motor networks. **B1.** Independent components showing activity modulations around error trials in the pMFC (top: note the increase in activity on the error trial) and occipital areas responsible for perceptual processing (task-relevant color encoding areas, on the bottom left, and task-irrelevant motion features, on the bottom right). Note the increased activity in post-error trials in relevant color areas and decreased activity in irrelevant motion areas. **B2.** Dorsolateral prefrontal and frontopolar increases (top) and left and right motor decreases (bottom) following error responses (adapted from Danielmeier et al., 2011).

### 1.4.3 Neural oscillatory mechanisms underlying error-monitoring and control processes

Neuroimaging studies have been influential to show the influence of the pMFC monitoring conflictive and error events and signalling other brain regions to regulate performance. Yet, the exact nature of the communication between the neural networks involved in action monitoring and control regulation only recently have been uncovered. In this regard, the study of *neural oscillations*[6] have shed new light in this matter revealing EEG dynamics that may reflect specific physiological mechanisms involved in the organization and communication of neural computations (Buzsaki & Draguhn, 2004; Fries, 2005; Varela, Lachaux, Rodriguez, & Martinerie, 2001), providing effective measures to understand the temporal coordination of information across brain regions, such as the imperative communication between monitoring and control networks.

During the last years of research, a growing body of studies showed that theta oscillatory activity (4–8 Hz) underlie the activation of a generic system responsible for the evaluation of demanding situations requiring greater cognitive control (Cavanagh, Zambrano-Vasquez, & Allen, 2012; Luu, Tucker, Derryberry, Reed, & Poulsen, 2003; Luu, Tucker, & Makeig, 2004).

The first evidence of error-related midfrontal theta activity was revealed by Luu and Tucker (Luu & Tucker, 2001). In their study they applied a high-pass filter (4-12 Hz) in the EEG data in order to filter out the large slow waves in which the ERN develops (P3-like components) and isolate an ERP signature specially related with error detection (see an example of this method in Figure 7a). This analysis was critical by showing that the ERN raises from frontal oscillations in the theta range. Other studies have been consistently showed that performance errors, negative feedbacks, response conflict and events involving uncertainty about actions and outcomes lead to an increase of medial-frontal theta power[7] (Cavanagh et al., 2009; Cavanagh, Frank, Klein, & Allen, 2010; Luu et al., 2003; Nigburg, Cohen, Ridderinkhof, & Sturmer, 2012; Pastotter, Dreisbach, & Bauml, 2013; van Driel, Ridderinkhof, & Cohen, 2012) (Fig. 7b). Moreover, mid-frontal ERP components such as the ERN, FRN or the N2 are parsimoniously characterized as reflections of theta band activities (Cavanagh et al., 2012).

Interestingly, it has also been shown that theta-band phase synchrony (which measures power-independent consistency of phase values across the two recording

---

[6] Neural oscillations represent rhythmic patterns of neural activity that can be driven by mechanisms localized within individual neurons or by large assemblies of neurons

[7] The most common measures are analysis of *Power* of the oscillatory activity (i.e. the magnitude of activation in a group of neurons) and *Phase Coherence* (Synchronization of oscillatory phases across neural networks). Phase Coherence measures can tag local communication between neurons within the same network and be used to examine functional connectivity between groups of neurons in large networks (see Varela et al., 2001).

electrode sites during a certain period of time) between mid-frontal and distal sites (lateral frontal; motor; sensory) is a common neural mechanisms underlying events requiring greater control processes (Cavanagh et al., 2009; Cavanagh et al., 2010; Narayanan et al., 2013; Nigburg et al., 2012; van Drial et al., 2012; van Vijver et al., 2011). These findings suggest that medial-frontal networks interact with networks involved in attentional control, motor and sensory processing by theta-phase dynamics. For example, Cavanagh et al (2009) showed that an amplified oscillatory synchrony in the theta band between mid-frontal and lateral frontal recording sites after errors. These findings are fascinating since it has been proposed that synchronization of neural oscillations may be important mechanisms by which long distant neural populations within a network can communicate (Fries, 2005). Moreover, in the aforementioned study (Cavanagh et al., 2009) it was observed that both the degree of oscillatory synchrony (i.e. functional communication between pMFC and LPFC) and the increase of theta power predicted the amount of PES.

Drawing in these evidences, pMFC theta activity stands as a plausible neurobiological mechanism by which the pMFC coordinates local and long-range neural networks to monitor actions and detect response errors.

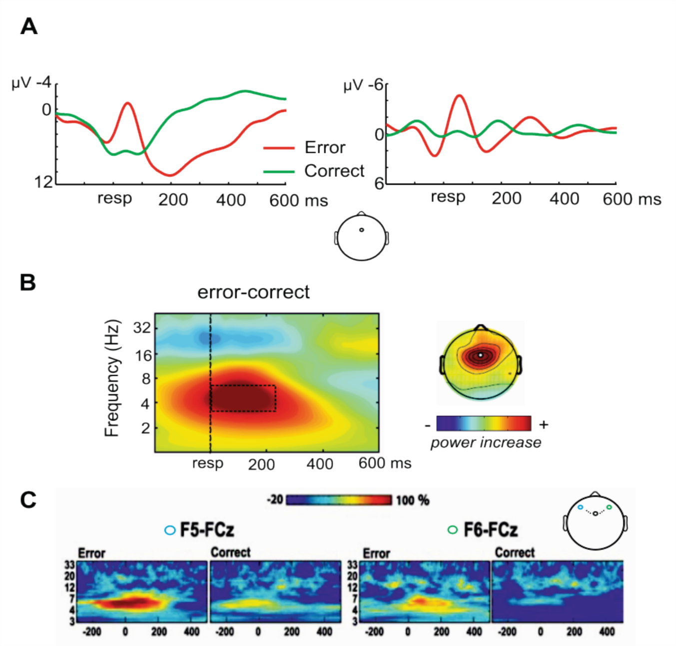


**Figure 7. (A)** R-Locked ERPs for correct and error responses showing the ERN in error trials. On the left side ERPs filtered with a low-pass filtered < 14 Hz. On the right side ERPs filtered with a band-pass filter in the theta range (3-9 Hz) (simulated data). **(B)** Medial-frontal theta power increase in errors and related topographical map (adapted from van Drial et al., 2012). **(C)** Phase synchrony between the Fcz electrode (black circle) and the fronto-lateral recording sites, F5(blue circle) and F6 (green circle), calculated as a percentage change from baseline. Note the increase of phase synchrony between the distant electrode locations in the theta band independent of the power amplitude in error trials (adapted from Cavanagh et al., 2009).

## 1.5 Theoretical and integrative accounts of error-monitoring

During this last 20 years of research a number of theoretical models, rooted in cognitive neuroscience, have arisen to explain the functional significance of the neurophysiological markers related to error monitoring and compensation processes, namely the ERN/FRN/N2 signals and action-monitoring related fMRI activations of the pMFC. In this section of the manuscript three influential theoretical accounts of error-monitoring and cognitive control would be briefly explained and some data supporting them will be described.

### 1.5.1 The Comparator Model: error-monitoring based on efference copies of the motor output

The *comparator model* proposes that the cognitive system may hold representations that an error as occurred or is about of being made, such that error-monitoring involves a comparator process in which the representation of the intended, correct response is compared to the representation of the actual response. In this framework the ERN reflects the output of the system whenever it detects a mismatch between the intended (correct) response and the actual erroneous response, as determined by the state of the response system after the response is executed (Coles, Scheffers, & Holroyd, 2001; Falkeinstein et al., 1990; Falkeinstein, Hoormann, Christ, & Hohnsbein, 2000; Gehring et al., 1993; Scheffers & Coles, 2000).

Because the ERN is an incredible fast signal it was first proposed that the comparison process could not rely on sensory or proprioceptive information, since this information is too slow and would not be available until the response has been completed; but instead would be guided by internal predictive (forward) mechanisms. Pioneering work by Cooke and Diggles (1984) first speculate on the possibility that fast error-corrective movements (e.g. when reaching a cup without looking at it) cannot rely entirely on proprioceptive external feedback of the motor output, but depend on predictive models of the motor plan (see also Angel, 1976), such that an error detector system may be fast enough to monitor when a wrong selection of the motor command is produced, enabling fast compensatory mechanisms (e.g. correcting the position and velocity of the hand while reaching the cup). For the comparator model the ERN represents an electrophysiological marker of an equivalent type of error detection system (or mechanism) operating, most likely, in the ACC and surrounding regions of the pMFC (Fig. 8).

In accordance with forward models of motor control, when a motor command is triggered the system generate a representation, or a prediction, of the potential consequences of ones actions, using an internal copy or emulator of the on-going motor command, i.e., *efference copy*, (von Holst, & Mittelstaedt, 1950) or a *corollary discharge* (Sperry, 1950). This motor copy interacts with the sensory processing system at several levels of the nervous system, being further used by the system to anticipate and, potentially, to cancel out any disturbance arising from conflictive sensory effects of movement (Crapse & Sommer, 2008; Desmurguet & Grafton, 2000). Studies in humans have been powerful in showing the influence of forward (predictive) models during motor planning, control, and learning (Desmurget & Grafton, 2000; Wolpert, Ghahramani, & Jordan, 1995; Wolpert & Miall, 1996; Wolpert, Diedrichsen, & Flanagan, 2011). Fascinatingly, it has also been shown that such predictive mechanisms have evolved in the animal kingdom too (*for a review see* Crapse & Sommer, 2008); without such adaptive mechanisms, moving animals would be impotent to determine whether disturbances registered by its sensory receptors would reflect real changes in the environment or simply a consequence of its own movement.

The comparator model suggests that during movement execution if an internal error signal is triggered, the system may implement fast inhibitory commands and error correction processes in order to prevent an error of being executed. In line with this idea, recently Burle et al (2008) have shown that the duration of the ERN is longer and the peak higher when errors are corrected. In this study it was observed that the ERN arises in the interval between erroneous muscle activations, or partial errors (recorded with EMG), and the onset of the correct muscle activation (i.e. error-correction). Others have proposed that the ERN may be elicited in parallel or immediately after the implementation of error correction processes (Rodriguez-Fornells et al., 2002). In line, it is also plausible that the ERN represents an alarm signal aimed for suppress proponent erroneous actions (Hochman, Orr, & Gehring, 2014), or it's involved in some sort of error-correction process, even before proprioceptive or sensory feedback is available.

A somewhat counterintuitive proposition of the comparator model is that error detection involves an explicit comparison between the executed response and a separate representation of the intended (correct) response. Given the fast onset of the ERN response and findings demonstrating error-related activity during partial (unconscious) motor error activations, it is unlikely that this neural signal represent an explicit error detection process but rather an automatic monitoring mechanism, dissociable from error awareness (Endrass et al., 2005; Nieuwenhuis et al., 2001), preceding the explicit perception of error commission (Yeung et al., 2004).

The model further states that after actual (overt) errors that the cognitive system uses internal error signals to implement remedial actions (Coles et al., 2001), as increasing the accuracy of the following responses or slow down the reaction time on ensuing trials following an error (Gehring et al., 1993).

Recently, two studies examining error-monitoring processes during skilled performances on piano had given support for a model of error-monitoring based on forward, predictive mechanisms (Maidhof et al, 2010; Ruiz et al., 2009). In these studies it was reported a negative component similar to the ERN just before the onset erroneous movement (i.e. playing an incorrect note on the piano - pitch errors), which was related to post-error slowing processes. The authors interpreted this early ERN as a neural error prediction signal at the level of motor preparation, which allow musicians to anticipate several notes in advance and detect upcoming errors even before the action is initiated or auditory feedback available.

Although the comparator model offers a simplistic account to understand the basis of error monitoring during motor movement, some criticisms may be signaled. First, the model lacks of neuroimaging evidences regarding the localization of the comparator (detector) error system. Further, the theory has not formalized an explanation for the modulation of the pMFC during response conflict, nor has explained the N2 component during correct conflicting responses; however, it is possible that the N2 can arise from subthreshold muscle twiches, i.e. partial errors', representing a mismatch between the intended and actual occurrence of inappropriate motor commands on trials with higher conflict. Finally, the model cannot account for findings regarding error-related brain activity associated with the monitoring of error feedback information.

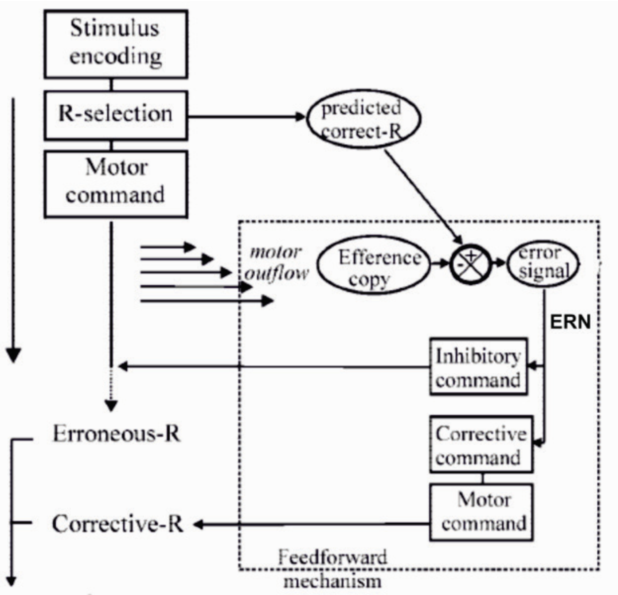


**Figure 8:** Schematic diagram of the different stages of information processing (*boxes*) and internal representations (*ellipses*) when a motor error is produced. The *dotted square* represents a feedforward (predictive) control mechanism. Accordingly, when an internal error signal is triggered, resulting from a mismatch between the predicted intended response and the actual state of the system (and giving rise to the neural marker ERN), the system may implement fast inhibitory commands or correction processes in order to suppress an error of being executed. However, if an error occurs the system is prompted to implement remedial actions, as for example PES (adapted from Rodriguez-Fornells, 2002).

### 1.5.2 The Conflict-monitoring Theory: errors as a form of response conflict

During the last years the *conflict monitoring theory* became one of the most successful and influential approaches to understand performance monitoring and cognitive control. Since its original formulation (Carter et al., 1998), a large body of experimental evidence has been accrued in support of the theory (Botvinick et al., 1999; Botvinick et al., 2001; Braver et al., 2000; Kerns et al., 2004; MacDonnald et al., 2000; van Veen & Carter, 2002; Yeung et al., 2004).

The theory was grounded on the hypothesis that during the performance of non-routine and challenging tasks cognitive control is recruited, in part, by a mechanism or a system that detects conflicts in information processing. Based in series of elegant neuroimaging studies the model offered an integrative account that accommodates a set of different findings showing that the pMFC, and specifically the ACC, is especially engaged in contexts requiring the overriding of prepotent responses, response selection under underdetermined or uncertainty situations, and performance errors (Botvinick et al., 2001; Botvinick et al., 2004; Carter & van Veen, 2007).

The first key point of the model is that the ACC is responsible for the monitoring of conflictive responses. Accordingly, response conflict occurs when prepotent but inappropriate response tendencies are elicited. In this framework errors are interpreted as special form of response conflict that occurs when incorrect response tendencies are not successfully overridden. In many cases, however, inappropriate response tendencies are suppressed in time, but accompanied by greater levels of conflict before the actual correct response is ensued. Accordingly, both errors and highly conflictive responses led to greater ACC activity, a prediction already confirmed by several neuroimaging studies (e.g. Carter et al., 1998; Kerns et al., 2004; Ridderinkhof et al., 2004).

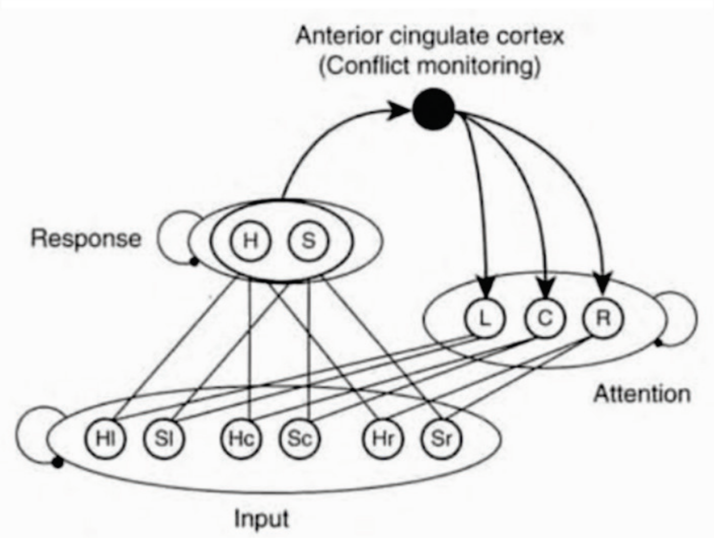


**Figure 9.** Conflict-monitoring model for the Eriksen flanker task. Stimulus processing units (the flanking letters, congruent or incongruent with the target) activate a given response. Conflict is generated in the response layer by the coactivation of mutually inhibitory response channels (i.e competing response activations). The ACC detects the presence of response conflict and signals the need of adjustments in a control network, as for example the increase of attention in order to reduce conflict from the flanking letters on subsequent trials (adapted from Botvinick et al. (2001)).

The authors offered an integrative account to explain fMRI results of pMFC activity and ERP findings showing the frontocentral ERN component after error commission (Falkeinstein et al., 1990; Gehring et al., 1993). In conformity with the model, errors take place when representations of more than one response are co-activated but the representation of the incorrect or undesired motor command is more quickly activated and reaches a threshold to initiate the actual response. Behavioral studies, for instance, show that errors are more likely in conditions of response conflict and most of the times result from premature responses (e.g. Gratton et al., 1988) probably caused by fluctuations in attentional control (Botvinck et al., 2001; Eichele et al., 2008). For the conflict model the ERN reflects conflict that develops in the period following error responses as a consequence of continued processing of the motor representations of the intended correct and actual incorrect response, such that conflict arises when there is an overlap between competing response activations (Fencick & Gehring et al., 2001; Botvinick et al., 2001; Yeung et al., 2004;). However, Burle et al., (2008) have put this assumption in evidence by showing that the amplitude of the ERN, in fact, decreases as the temporal overlap between the two response activations decreases.

Moreover, the theory offers an explanation for the observation enhanced amplitude of the frontocentral N2 component in interference tasks as ERP marker of pMFC activity before highly conflictive correct responses are initiated (van Veen & Carter, 2002; Yeung et al., 2004). In this sense, the theory offers an integrative account of the N2 and the ERN as common underlying neural mechanism.

The theory further suggests that upon the detection of conflicting response tendencies the ACC signals other areas of the brain-related with a control network, such as the lPFC, to implement strategic adjustments in cognitive control (MacDonnald et al., 2000; Kerns et al., 2004), which serve to prevent conflict in subsequent performance. In this sense, the model accommodates very well behavioral findings in reaction-time tasks showing that RT costs and error rates in conflicting events that are immediately preceded by conflicting trials are reduced (e.g., Gratton, Coles, & Donchin, 1992; Kerns et al., 2004). Several studies provided empirical support for this hypothesis by showing, first, that in high-conflict trials the magnitude of pMFC activity predicts the degree of behavioral adjustment and the activity level in the lPFC on the subsequent trial (Kerns et al., 2004; Egner & Hirch, 2005).

Despite on-going debates the conflict monitoring theory has been influential and has accommodated a large body of neuroimaging and behavioral findings in force-choice reaction time tasks involving competition between competing motor plans and contexts of increased task difficulty. However, some questions remain to be addressed. An important criticism to the conflict-monitoring model is the lack of evidences from animal studies confirming the presence of neurons in the ACC that signal specifically the presence of response conflict (Ito et al., 2003). Furthermore, the response-conflict model has also remained silent regarding feedback-related activity in the pMFC in presence of negative feedbacks as well as the presence of enhanced medial-frontal theta activity during events involving increased response competition and response errors.

The extent to which the conflict-monitoring system is also sensitive to the presence of conflicting events that cannot be consciously reported, or to non-motor forms of conflict such as emotions, ideas or thoughts, which many times are effortful, are questions that so far have not been addressed.

Finally, although the conflict-monitoring model has been developed upon neuroimaging evidences showing larger activations of the pMFC on more challenging situations or non-routinary situations (e.g. when for example one is learning a new skill), no predictions had ever been made on how the conflict-monitoring system is modulated during procedural or associative learning.

### 1.5.3 The Reinforcement learning Theory: prediction error signals

The Reinforcement Learning theory (RL) of error processing (Holroyd & Coles, 2002; Nieuwenhuis, Holroyd, Mol, & Coles, 2004) provides a framework for understanding the mechanisms of error-monitoring according to principles of reinforcement learning in humans (Sutton & Barto, 1998).

The theory is founded on previous research that implicates the basal ganglia and the midbrain dopaminergic system in the computation of reward prediction signals during learning (Schultz, Dayan, & Montague, 1997; Schultz, 2007). Accordingly, the basal ganglia monitor ongoing events and compute predictions on whether the outcomes associated with those events are favorably or unfavorably. From this revision a prediction error signal is sent to the system informing whether the outcome is better (positive prediction error) or worse (negative prediction error) than expected. These negative and positive error signals are conveyed from the basal ganglia as phasic decreases and increases, respectively, of activity of the mesencephalic dopaminergic system. Crucial for the flexibility and adaptability of the error system, these phasic signals are used by the basal ganglia to update its predictions and then are sent to the frontal cortex, targeting the pMFC. The pMFC, as a hub of action monitoring and selection, would then process this error information and select appropriate behaviors to guide learning and assist goal-adaptive behavior (Holroyd & Coles, 2002; Holroyd et al., 2004; Rushworth, Walton, Kennerley, & Bannerman, 2004).

An important assumption of the RL theory is that the pMFC acts as “control motor filter” deciding in a given context which motor commands are the more effective based on the reinforcement learning signals conveyed by the dopaminergic system. In this view, the pMFC enables distinct neural “control” structures (e.g. DLPFC, amygdala, orbitalfrontal cortex) that project to the pMFC to take command of the motor system, by filtering or deciding which controller is best suited to address the task at hand. In agreement, it has been demonstrated that the pMFC has a critical role on the integration of action-outcome associations across trials and on the computation of the subjective values of actions (Jocham, Neumann, Klein, Danielmeier, & Ullsperger, 2009; Matsumoto, Matsumoto, Abe, & Tanaka, 2007; Walton, Devlin, & Rushworth, 2004).

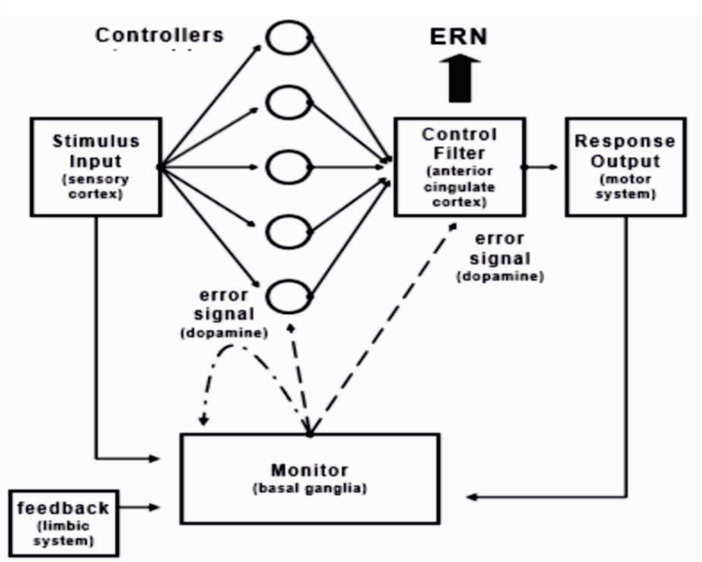


**Figure 10.** Multiple controllers (e.g. DLPFC) in the brain produce actions based on sensory information. All these commands are filtered by the ACC (control filter) such that only appropriate and non-conflicting responses pass to the motor system. A monitor located in the BG processes feedback information (such as rewards or punishments) and efference representations of ongoing motor commands. The monitor computes predictions on whether these feedback information is favorably or unfavorably, producing error signals that are conveyed by the dopaminergic system to other areas, including the ACC (main target), where they reinforce adaptive and goal-directed processes. The ERN is modulated by the impact of the error signals on the controller (adapted from Holroyd et al., 2004).

The theory also offers an integrative account for the observation that the pMFC is activated by error information, regardless of whether the source of information is internal (i.e. available from the motor effector) or external (feedback-dependent). Based on a neuro-computational model, the theory has explained the striking similarity of ERP components following error responses (ERN) and error feedbacks (FRN) (Miltner et al., 1997; Holroyd & Coles, 2002) as the same signature of error prediction. In this framework both response and feedback ERN components reflect a reinforcement learning signal that is transmitted to the pMFC, such that the pMFC does not itself monitor errors, but rather reflects the output of a system indicating whether outcomes of actions are better or worse than expected. In electrophysiological terms, the ERN signal results from the disinhibiting of the apical dendrites of motor neurons in pMFC when a negative prediction error signal is sent from the basal ganglia to the pMFC. However, recent findings have challenged the view that the ERN may actually represent a neural signal selectively involved in the evaluation of poor (or unfavorable) performance, but rather is modulated by the degree of discrepancy between expected and real outcomes, that is, it reflects expectancy violations (Cavanagh et al., 2011; Mas-Herrero & Marco-Pallares, 2014; Oliveira et al., 2007).

Nevertheless, the theory predicts that the response or feedback ERN is elicited following the earliest predictor of a negative, unfavourable (or unexpected outcome). To test this hypothesis the authors used a probabilistic learning task with 2-alternative response with either random (*unlearnable*) or fixed (*learnable*) stimulus-response mappings to examine the relationship between feedback and error related brain activity, respectively, as learning progressed (Holroyd & Coles, 2002; Holroyd et al., 2005). The reasoning behind this approach was that it would be only possible for subjects to decide whether a given answer is correct without relying on feedback in the condition of fixed stimulus-response mappings, then the first indication of an error would be indexed by the ERN. On the other hand, in a novel environment, in which stimulus-response mapping needs to be learned, error information is not available until the delivery of an external error feedback, leading to the FRN appearance. As predicted the authors found that during learning conditions (or during early learning stages) the ERN was larger after feedbacks (eliciting an FRN) and diminished during the response moment (ERN). However, after subjects have learned the stimulus-response associations the pattern reversed and the response ERN predict the value of the response. Moreover, fMRI results show that both kinds of errors signals (response-based and feedback-based) activate the same structure, that is, the dorsal ACC, but with a different temporal dynamic (Holroyd et al., 2004; Mars et al., 2005).

Although the RL theory of error processing has brought interesting and testable ideas to the research field regarding action monitoring and cognitive control, there are some difficulties to accommodate the neurochemical mechanisms underlying the RL theory in the laboratory context of error processing. In first place, some of the predictions of the RL theory are only possible to test using invasive methodology: intracortical recording, pharmacological studies or animal studies; However, not many studies have been performed until the current moment (see an interesting opinion on this matter on a review paper by Ullsperger et al., 2014). Another delicate issue regarding the theory is the direct assumption that the ERN reflects a decreased of phasic DA neuron activity in the BG and related projection areas (fundamentally the ACC), as some authors have suggested that the midbrain dopaminergic system lacks the temporally precision to generate a fast signal error prediction signal as the ERN (Jocham & Ullsperger, 2009). Moreover, it’s not very explicit by the theory how conflict is treated according to the reinforcement dopaminergic signals. Finally, no much evidence has been acquired supporting the existence of reinforcement learning signals during procedural or motor learning (but see a comment on Wolpert et al., 2011), as the model has only been tested in experimental settings involving associative learning.

In sum, the RL theory offers an account for the role of the pMFC monitoring internal and external error information and action selection based on reward-related functions associated with midbrain dopaminergic system.

## 1.6 Towards an ecological approach to study the electrophysiological mechanisms of human error-monitoring

So far I have reviewed the most relevant behavioral and neurophysiological findings regarding the mechanisms underlying error-monitoring and the consequent implementation of adaptive regulatory behaviors. Those studies, up to now, offer an extensive and detailed picture of the current state of art and constitute the ground of influent theories regarding performance monitoring and cognitive control processes.

Important, however, it is worth mentioning that the majority these research efforts have been mainly focused on the observation of isolated, low-level, performance errors in fairly simple tasks, most of them using force-choice reaction time procedures. As commented before, in the standard laboratory paradigms, such as the Stroop or the Erikson Flanker tasks, errors most commonly exemplify *action slips* or *lapses*, i.e. actions that are carried out automatically without conscious deliberation, resulting in an "unintended error of execution of a correctly intended action" (see Norman, 1981; Reason, 1990). In these contexts, action slips, most of the times, result from perceptual lapse, attentional fluctuations, or maladaptive control mechanisms.

However, in the real world errors can have different forms[8] and be very different in nature, depending on the contexts in which the individual is inserted. In this sense, not all errors may result from problems of activation of attentional or control mechanisms of the nervous system. In fact, errors can still take place in conditions whereby the amount of attentional control deployed to the task on-hand is appropriate. For instance, when learning a new skill or procedure, whereby representations of the task to be performed are week and the subjective evaluation of on-going performance is difficult and underdetermined; when errors are due exogenous sources in the environment (machine malfunctions; accidents caused by a third party). Yet, the rapid implementation of flexible cognitive control resources may be beneficial to overcome potential hazards from these classes of errors.

---

[8] Seminal work by James Reason (1990) and Donald Norman (1988) has been influential to categorize different type of error forms. A particularly interesting taxonomy dissociates slips of action from mistakes.
Slips make reference to unintended actions or execution failures (e.g. taking a wrong turn while driving).
On the other hand, mistakes are more related with planning failures or errors in the formation of an intention, and therefore may result from failures in higher-level processes involving decision-making and learning. For example, if a person makes a mistake in choosing a goal, the action chosen to reach that goal may be corrected or acceptable, but since the goal is incorrect or misadjusted the outcome may not be as desired

Furthermore, there are several clinic conditions of individuals showing abnormalities in the control of their actions which itself are related to the occurrence of unintended errors of action (e.g. Parkinson disease, obsessive-compulsive disorder, utilization behavior, optic ataxia, alien-hand syndrome, delusions of control in psychotic conditions, and many others). On top of that, humans when interacting with complex machines or systems also display different tolerance to error feedback information and, therefore, to cope with this information they may be bias to commit different type of deliberate or 'intentional' erroneous actions[9] which may reflect the externalization of distinctive cognitive schemas and standards of self-reinforcement to deal with uncertainty or stress (Bandura, 1989).

Therefore, an important limitation of most of the previous studies using reaction-time tasks, which have mainly focused on the examination of slips of action, is that they do not entirely cover the complexity of real environments or novel and learning situations in which human agents have been found to commit errors (Reason, 1990; Sanders & Moray, 1991; Woods, 1994). Yet, as Norman commented on his seminal paper "Categorization of Action Slips" (1981), a detailed analysis of action slips can be beneficial to construct the outlines of a theory of action that explains how an intention is represented and acted upon. In a certain way Norman was right. In fact, despite their simplicity, these tasks have been suitable to study different aspects of error processing, such as, pre-response conditions of conflict, the error response itself and post-response processes involving error detection and regulation. As these processes, to a certain extent, are likely to be present in complex real tasks, the understanding of their functional characteristics, their temporal properties and their neural underpinnings is certainly valuable to comprehend the cognitive and neural intricacies of human fallibility out of the laboratory settings. The studies I have reviewed in the last sections seem to support this approach.

Still it is important to move forward this research including more ecologically valid contexts in order to explore in more depth the neural mechanisms, and its related mental processes, underlying action monitoring. At this time point, the field is lacking the implementation of novel and creative experimental paradigms combining complex and multimodal tasks, which involve extended interactions between the performer and the environment and the expression of complex intentions of the performer, with the available new neuroimaging techniques.

It is worth mentioning, though, that recently some efforts have been made in this direction and some researchers have started to explore error-monitoring and compensatory processes in more complex activities such as typewriting (Logan & Crump, 2010), music performance (Ruiz et al., 2009; Maidhoff et al., 2010), shooting (Bediou et al., 2012), tasks involving visual searching and exploration (Solman,

---

[9] By intentional erroneous actions I make reference to errors that are due to planning failures or mistaken behavior.

Cheyne, & Smilek, 2012), motor tracking tasks, or driving in realistic car simulators (Zhang et al., 2013). However, these studies represent only a small and timid attempt in this promising research field; and not all of them have adventured in examining the neural mechanisms underlying error monitoring in such complex contexts. Yet these brave and creative approaches may inspire and encourage new directions for future research.

The present thesis has as main motivation follow the steps given by these aforementioned studies, standing as an intention to bring fresh and renovated air to the on-going research. In this work I will mainly direct my attention to electrophysiological responses, which have been already well described in the literature of error processing, in new experimental settings and introducing new concepts and questions that all over these years have remained elusive, difficult to address or have not even been considered. I believe that the questions that I lay out in what follows may extend and deepen our current knowledge on the flexibility of the error monitoring mechanisms sustained by the human brain, opening new avenues for future research.

1. To what extent is the pMFC error-monitoring system involved in the monitoring and regulation of unattended conflicting/error-prone events that cannot be consciously reported due to limited attentional resources?
2. How practice leads to functional changes in the brain mechanisms associated with error-monitoring during the acquisition of a new motor skill, as for example learning to play a new instrument or a musical rhythm? And how does our subjective perception of error-awareness evolves during this learning process?
3. How do we recognize ourselves as the agents of our actions? Do we use the same error-monitoring mechanisms to distinguish self-generated errors from errors that are not intentionally caused by ourselves but still carried out by our own body?
4. Do humans have the same tolerance to erroneous feedback information? How individual differences in error tolerance, for instance, may explain different type of mistakes performed by human agents in their interaction with real-working and error-prone contexts? To what extent are these individual differences reflected in neurophysiological mechanisms indexing specific cognitive and motivational states underlying external feedback monitoring.

In the four sections that follow, I will introduce the reader with the theoretical framework in which these four general questions that constitute the core of the present thesis were developed and afterwards, in chapter 2, I will present the Research Aims (Objectives) of the present Thesis dissertation.

### 1.6.1 Monitoring and regulation of exogenous unattended sources of conflict

A long lasting debate in cognitive psychology and neuroscience concerns the way unattended information in the environment is processed and influence goal-directed behavior (Allport, Tipper, & Chmiel, 1985; Broadbent, 1958; Broadbent, 1982; Hillyard, Vogel, & Luck, 1998; Lachter, Forster, & Ruthruff, 2004; Miller, 1991; Posner, 1994).

Attentional selection is needed to protect the limited capacity of the decision-making apparatus from information overload. Still there are many salient, but unattended, visual inputs in the environment to which we do not have an explicit conscious perception[10] and which may be automatically processed influencing our ongoing behavior. To illustrate, imagine yourself driving your car back home and suddenly a dog on the sidewalk starts running in the direction of the street. With most of your attention deployed to the oncoming traffic you may neglect the dog's presence; but still you may find yourself automatically reducing the velocity and correct the car trajectory overcoming the potential hazard.

For long time many authors have suggested that cognitive control functions, such as the detection of conflicting and error-prone events (the dog running in direction of the street) and the consequent implementation of inhibitory or regulative behaviours (reduce the speed), are effortful and regulated by attentional (top-down) control processes of prefrontal neural networks (Dehaene & Naccache, 2001; Jack & Shallice, 2001; Posner, 1994; Posner & DiGirolamo, 1998). Nevertheless the example given here, along with other behavioral findings (Miller 1991), suggests that selective attention cannot eliminate completely all unattended information from further processing and that relevant, but unattended, visual inputs still influence our behavior and, hypothetically, brain networks related with 'high-level' cognitive control functions so that appropriate courses of action can be initiated in face of conflicting or error-prone situations.

Furthermore, recent findings have shown that 'high-level' control mechanisms of the prefrontal cortex, such as error-monitoring (Hester et al., 2005; Nieuwenhuis, et al., 2001), inhibitory control (Sumner et al., 2007; van Gaal, Ridderinkhof, Fahrenfort, Scholte, & Lamme, 2008) or task set preparation (Lau & Passingham, 2007) can unfold under conditions of limited conscious perception and explicit knowledge.

In the first study of this thesis (Chapter 3) we will investigate the extent to which prefrontal neural networks related to action monitoring and regulation are responsive to the presence of unattended, yet salient and error-prone, events that cannot be consciously reported due to limited attentional resources. Further we examine which neural oscillatory mechanisms orchestrate the flexible monitoring

---

[10] Here I refer to conscious perception as the explicit knowledge about the visual representation of a given object, which can be further used for verbal report, also referred to as access consciousness (see Lamme, 2003; Block, 2005)

and compensation of inappropriate action sets that are automatically triggered by unattended visual stimuli that cannot be consciously accessed and acknowledge.

To address this question we took advantage of the high temporal resolution of the EEG while subjects performed a novel variant of the Erikson flanker task, in which spatial attention was manipulated preventing the conscious access of unattended visual stimuli presented parafoveally. Previous research has highlighted the importance of spatial selective attention in conscious perception such that visual stimuli that are not attended usually cannot be consciously available for verbal report (Cohen, Cavanagh, Chun, & Nakayama, 2012; Lamme, 2003). Therefore, with this novel redesign of the Flanker task we aimed to simulate a complex visual context in which participants selectively prioritized visual information that was more relevant to the task on-hand at the expense of other items that, although visible, remained oblivious. This experimental manipulation allowed us to study behavioral indexes and electrophysiological markers of MFC activity related with the monitoring and regulation of both attended and unattended conflicting visual events that are likely to induce erroneous actions and which require rapid behavioral adjustments.

### 1.6.2 Error-monitoring and error-awareness processes during the acquisition of motor skills

Error is a mechanism integral to learning, one necessary for developing skills in complex tasks. Without the ability to detect and use information from errors, it would be impossible to develop simple adaptive behavior and to achieve high levels of motor skill as that found in top-level musicians or elite athletes. However, understanding how an individual learns from its own errors and progresses to excelling levels of performance has not been a trivial process. Theories on motor skill acquisition have for long time emphasized the importance of error-detection processes (Adams, 1971; Schmidt, 1975; Wolpert & Kawato, 1998) and the interest regarding the neural circuitry underlying motor skill learning has proliferated (Hikosaka et al, 2002; Ungerleider et al., 2002). In this regard, a large body of neuroimaging studies has shown the critical participation of medial-frontal structures of the brain, and striatal regions, during early stages of skill acquisition (Floyer-Lea & Matthews, 2005; Jenkins, Brooks, Nixon, Frackowiak, & Passingham, 1994; Jueptner et al., 1997; Petersen, van Mier, Fiez, & Raichle, 1998; Ramnani & Passingham, 2001; Toni, Krams, Turner, & Passingham, 1998). The involvement of these structures in early stages of learning is likely to reflect high level of monitoring of inappropriate response tendencies or competition between multiple motor plans, and, consequently, signaling the need for increased control processes during stages of learning wherein performance is more demanding (Botvinick et al., 2001; Jueptner et al., 1997; Paus, Koski, Caramanos, & Westbury, 1998). However, the extent to which the learning of a new motor skill leads to electrophysiological changes related to error-monitoring and error-awareness has not been investigated.

To address this question in the study 2 of the present thesis (Chapter 4) we designed a task requiring the learning of series of musical rhythmic patterns and tracked changes in ERP signals previously linked to error-monitoring and error-awareness processes across distinct learning stages. We used ERPs because, unlike other neuroimaging approaches, it is a technique particularly suited to investigating the time course of fast cognitive processes, like error detection processing in a context involving music-like performance (e.g. Ruiz et al., 2009).

The exploration of error processes during the process of rhythm production learning is also very interesting because of the type of errors that are observed. The type of errors observed during skill learning differs from the type of errors commonly observed in standard reaction time tasks in which errors are discrete and binary (i.e. present or absent). In rhythm learning, for instance, errors are more relative and therefore can range in magnitude (the degree of asynchrony of the executed motor action and the timing of the associated note) and change over the time. Moreover, during rhythm learning the subjective experience of error commission leading to error-awareness relies on internal evaluations of the performer which depend on the consolidation of internal representations or "memory motor traces" (Adams et al., 1971) that are refined throughout training (Wolpert et al., 2011).

Because the production of musical rhythms makes same unique demands for the executive and motor systems of the human brain, the study presented in Chapter 4 is hopeful to understand how brain systems involved in error processing and cognitive control are engaged in an ecologic context that mirrors many situations in which we acquired new skills.

### 1.6.3 Monitoring self-generated errors and agency violations

Whilst typing on my computer to write this manuscript, I reach for my coffee, take a sip, and replace the cup without even looking; automatically monitoring minor errors and implementing rapid adjustments on my movements, which are normally unavailable to awareness. This pragmatic and effective attribute of my motor system may provide me with the sensation of being causally responsible for reaching the coffee cup and, therefore, causally responsible of my own actions, including my own errors.

As for the most part we remain unconscious with regard to our *sense of agency*[11] (Gallagher, 2000) it may be remote or even absurd to think about the chance of experiencing our body "performing" actions that are not intended by us and which

---

[11] The term 'sense of agency' refers to the experience of controlling one's own actions, and, through them, events in the outside world.

may cause bizarre erroneous actions. Nevertheless, there are disparate clinical cases of individuals experiencing abnormalities in the control of their actions and related problems in their experience of agency. For instance, schizophrenic patients often report that their actions are not their own but are imposed by some other agent, these experiences are often labeled as delusions of control and hallucinations (Daprati et al., 1997; Frith, 2005; Synofzik, Thier, Leube, Schlotterbeck, & Lindner, 2010). Some other neurologic class of patients has been described because their hand often moves 'of its own accord', without patients will, known as the alien-hand syndrome (Marchetti & Della Salla, 1998).

So, a question that arises here then is how do we recognize ourselves as the agents of our actions? Is our brain able to distinguish self-generated errors from "alien-errors" that are performed without subjects' will and which may undermine one's natural sense of agency? Despite its importance, these fundamental questions in cognitive neuroscience have remained elusive and proven difficult to address experimentally.

In the study 3 (chapter 5) of this thesis we propose a creative and novel possibility to address these questions. Here, we put forward for consideration the existence of two dissociable neural *"error-monitoring systems*" which are responsible for the detection of self-generated errors and errors that represent agency violations, such as the pathological cases of "alien" actions.

Accordingly, when executing most of our actions an *internal error-monitoring system* is constantly in charge of detecting deviations from the expected goal of our on-going actions, even before the action is finished; and to implement fast error correction processes, or compensatory mechanisms if an erroneous tendency is detected (e.g. Gehring et al., 1993; Marco-Pallarés et al., 2008; Rabbit et al., 1966a; Rodriguez-Fornells et al., 2002). This incredible fast error-monitoring system, which may run outside of our consciousness, may be explained by forward predictive models of the motor command (Desmurget and Grafton, 2000; Wolpert et al., 1995; Mial & Wolpert, 1996), which allow error detection prior to sensory or proprioceptive feedback. As it was discussed in a previous section of this Thesis (*see* the comparator model hypothesis in section 1.5) the ERN, an ERP response appearing right after error commission, may stand as a possible neural index of this internal error-monitoring system (Gehring et al., 1993; Rodriguez-Fornells; Scheffers & Coles, 2000).

While this internal error-monitoring system may help the organism to adapt in a fast and flexible way to self-produced errors, another error-monitoring system may be responsible for register and evaluate if the sensory consequences following one's actions are in line with ones' usual experiences, providing us with the feeling of being the causal agents of our actions (Synofzik et al., 2008), or in other words, giving us a sense of control over our actions. Several models have proposed that during voluntary movements an internal (efference) copy of our motor actions is forwarded and transmitted to the sensorimotor cortex in order to provide information about the expected sensory consequences of our actions (Frith et al., 2000; Jeannerod, 2005; Wolpert & Miall, 1996). In this sense, the sense of control over our actions may ultimately rely on multimodal reaferent feedback (visual, auditory, proprioceptive and somatosensory inputs), such that if the comparison between the predicted forward model and the actual feedback is coherent (i.e. if the system detects no errors or discrepancy between both) we may fell as "I did that" (Haggard & Chambon, 2012; Synofzik, Vosgerau, & Newen, 2008).

In the study 3 (chapter 5) of this thesis we tried to prove a functionally dissociation between the two error-monitoring systems and to examine its related neural delays in processing by examining the extent to which both systems can be distinguished using neurophysiological (ERP) signatures.

To accomplish this goal an innovative experimental situation was invented in a virtual reality setting. Healthy participants were embodied into avatars and then carried out an error-prone reaction time task while fully immersed in the virtual environment. Critically in certain occasions participants were deceived regarding their actions, whereby their intended correct actions were violated, ending in errors of agency. This setting was aimed to simulate a possible, although remote, scenario in which the actions "I" try to implement in my body are falsified and influenced by an "alien agent". Then, we explored ERP signatures in two specific conditions: when the avatar was performing or introducing alien-errors and those situations in which the participant performed real-errors.

### 1.6.4 Individual differences in error-tolerance and decision-making processes

Common sense suggests that humans have different degrees of tolerance to errorful feedback information and that these differences may influence the way they make decisions during their interaction with error-prone and real-working environments. To illustrate consider the following example: during a flight simulation two pilots (A and B) were tested in a new procedure in which they have to learn specific rules of navigation by evaluating different types of external feedback which were informative about their performance. In the beginning of their training both were told that in some occasions they would receive positive feedback that would reinforce their current behavior while in other occasions they would receive negative feedback that could be either misleading or implying a new strategy and consequent changes on on-going behavior. Important, they were advised to change behaviour only when they were sure that a negative feedback required behavioural adjustments. The evaluators observed that both pilots committed different type of mistakes when they face negative evaluations. Pilot A compared to Pilot B, after learning a new rule (or procedure) was more prone to perseverate on the new learned rule for longer periods of time and more averse to shift his behaviour even when he received negative feedback signalling the need of changing the current strategy. In contrast, Pilot B, after receiving negative feedback, wish sometimes were misleading, changed his behavior almost automatically.

This fictitious scenario suggests that during the operation of complex technological systems involving rule-based decisions, humans show different degrees of tolerance to change pre-activated courses of actions based on erroneous feedback information. Therefore one may ask if (i) these differences are relevant to predict the pattern of mistakes performed by operators that intentionally formulate different action plans to carry on the same task; and if (ii) these differences may be explained by specific neurophysiologic mechanisms related to error feedback monitoring.

These questions are fundamental as the most devastating catastrophes that occur are somehow related to human decision processes in the operation of complex technological systems (e.g., drivers, air traffic controllers, etc.) (Casey, 2006; Wiegmann & Shappell, 2003). In this regard, it may be valuable to explore how individual differences in error-tolerance may affect human decision processes and the extent to which brain mechanisms involving the evaluation of error feedbacks may explain these differences. This assumption may gain some support by a large body of studies which have shown that individual differences in personality and temperament are related to different modes of self-regulation, showing variability in action-monitoring related EEG cortical activity (Foti & Hajcak, 2009; Luu, Collins, & Tucker, 2000; Padrao, Mallorqui, Cucurell, Marco-Pallares, & Rodriguez-Fornells, 2013).

However, using choice reaction time procedures to access these differences may be problematic because those tasks can only assess a particular type of errors, called action slips. Therefore, we may benefit from looking at different type of errors that somehow are due to individual differences related to the monitoring of complex chains of commands (Norman, 1988; Reason, 1990;) and which may echo different cognitive biases and motivational appetitive processes of self-reinforcement (Bandura, 1989)

A type of errors particularly suited to access those differences is the so-called rule-based mistakes (Reason, 1990; Rasmussen, 1983; Norman et al., 1988). Rule-based mistakes make reference to actions that match intentions but do not achieve their intended outcome due to incorrect application of a rule or inadequacy of the plan (usually, a rule that is frequently used, or it has been used with success in the past, and seems to fit the situation well enough, but in fact is inadequate). This type of mistakes may include errors in judgment, inference, and interpretation. Rule-based mistakes in a certain way are indirectly assessing the degree of error-tolerance of an individual

In the study 4 of the present thesis (Chapter 6) we examined how individual differences in error-tolerance determine the pattern of mistakes performed by humans in a learning context involving rule-based decisions that simulates the example of the two pilots during the flight simulation. Further it was investigated whether these differences could be reflected in neurophysiological mechanisms indexing specific cognitive and motivational states underlying external feedback monitoring.

# 2

# Research Aims

## Research Aims (Objectives)

This Thesis includes four studies aimed to explore distinct aspects of the electrophysiological mechanisms underlying human error-monitoring in different contexts. The questions behind those studies have been introduced in the last section preceding this chapter. As the answer to those questions is not trivial, in the present research we either designed novel experimental paradigms or recycled already known experimental tasks (for example the Flanker task) but approaching them from different perspectives. We applied electrophysiological measures, mainly event-related potentials (ERPs) and time-frequency analysis, to examine different neural signals related with error and feedback monitoring in these contexts. The goals of each one of those studies are briefly resumed below.

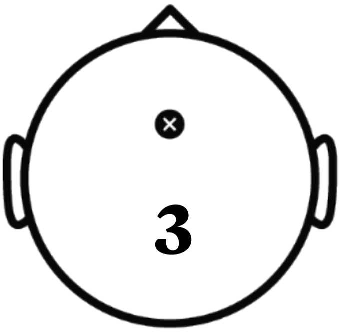
**In study 1,** we explored the possibility that the monitoring and compensation of unattended, but conflicting/error-prone, events which cannot be consciously reported due to limited attentional resources can still be mediated by the pMFC. In this study we recorded ERPs while participants performed a novel variant of the Erikson flanker task, in which spatial attention was manipulated preventing the conscious access of visual events presented at unattended locations. The manipulation of participants' spatial attention in this task allowed the examination of behavioral indexes and electrophysiological markers of pMFC activity related to the monitoring of both attended and unattended conflicting items. The goal was to investigate (i) how fast is the pMFC detecting error-prone/conflicting unattended events and implementing compensatory (reactive control) mechanisms, and (ii) which neural oscillatory mechanisms may underlie these processes. To accomplish this goal we combined ERPs, time-frequency and source localization analysis.

**Study 2** sought to explore the extent to which electrophysiological signatures related to error-monitoring and error-awareness changed during the learning of a new motor skill. To accomplish this goal we designed a novel experiment in which musically naïve individuals learned to reproduce a series of musical rhythmic patterns across 12 trials of learning, by synchronizing tapping movements with auditory events with different time (rhythmic) intervals. Then we examined changes in ERP signals across distinct learning stages in order to understand how practice leads to functional changes in the brain mechanisms associated with error-monitoring and error-awareness. It was hypothesized that early in learning, when internal representations of the rhythms were weak and greater demand for cognitive control processes was required to compensate erroneous/conflicting responses would lead to increased activity of the MFC as indexed by the ERN component. It was also expected that as learning progresses we would observe a decrease of error-related ERN activity reflecting less engagement of the error-monitoring system during more fluent performance. Furthermore, we predicted that the development of stronger internal representations of the target rhythms throughout the learning process will be related to greater emergence of error-awareness, indexed by the activity of later centroparietal positive components.

In **study 3** we investigated whether our feeling of agency would be the same when our body carries out actions or errors that were not intended by ourselves, but rather imposed by an ”alien” agent. The idea was to examine to what extent brain error-monitoring signals could distinguish self-generated errors from errors that are externally caused without one’s will.

We took advantage of current technological improvements in virtual reality and set up an intriguing experimental setting in which we recorded ERPs of healthy participants embodied into an avatar body while performing an error-prone reaction time task fully immersed in a virtual reality environment. Critically, in certain occasions participants were deceived regarding their own actions, wherein they observed their correct movements being falsified by an “erroneous” movement of their embodied virtual arm, violating their internal intentions and causing an alien-error. This setting allowed us to disambiguate the neural processes underlying the monitoring of self-generated errors and errors that are caused without subjects’ will and which may undermine one’s natural sense of agency. This study puts forward for consideration the participation of two dissociable neural *“error-monitoring systems”* during the processing of self-generated errors vs. errors that represent agency violations.

In **study 4**, we focused our attention on the concept of error-tolerance, i.e. the tendency of humans to follow well-programmed action plans (e.g. well-known routines) in the presence of conflicting or redundant error feedback information without the need to change pre-selected courses of action. The idea was to explore neurophysiological indexes of error-tolerance, with the purpose of characterize the cognitive and motivational aspects underlying different strategies adopted by human operators when coping with errorful feedback information in error-prone environments. To explore individual differences in error-tolerance we observed the performance of 80 healthy participants in a probabilistic reversal learning task and then compared the ERP responses of two groups of participants (with High and Low error-tolerance) which drastically differed in their propensity to maintain newly learned rules after receiving spurious negative feedbacks. We expected to observe ERP differences regarding the motivational engagement and attentional impact during error feedback anticipation and error feedback evaluation between individuals with Low and High error-tolerance.



# 3

# Study 1

## Exogenous capture of medial-frontal oscillatory mechanisms by unattended conflicting information[*]

---

[*] This study corresponds to:
Padrão, G., Rodriguez-Herreros, B., Zapata, L.P., Rodríguez-Fornells, A. Exogenous capture of medial-frontal oscillatory mechanisms by unattended conflicting information. (**submitted**)

## INTRODUCTION

Our senses are constantly overwhelmed with a huge amount of stimuli, being many of them not relevant to our current task goals. Thus in many cases most of these stimuli remain unattended, receiving very little processing beyond the analysis of basic physical properties (Broadbent, 1958; Lachter et al., 2004) (Broadbent, 1958; Lachter et al., 2004). However, there are many situations during our everyday life wherein we have to rapidly monitor relevant unattended information, which sometimes we do not have a clear conscious perception[1] so that appropriate behaviors can succeed. To illustrate this idea, imagine yourself driving your car back home and suddenly a dog on the sidewalk starts running in the direction of the street. With most of your attention deployed to the oncoming traffic you barely register the dog. However, even without being able of explicitly identifying the dog's presence, you find yourself automatically adjusting your current speed and trajectory.

There is now a wealth of evidence that the monitoring and compensation of erroneous and conflicting events is accomplished by a neural system sourced in the medial prefrontal cortex (mPFC) (Botvinick et al., 2001; Carter, 1998; Brown & Braver, 2005). EEG studies, for instance, have consistently described cortical responses with negative polarity in frontocentral electrodes of the scalp that indicate the activation of the mPFC during the occurrence of response conflict, the N2 component (van Veen & Carter, 2002; Yeung et al., 2004) and response errors, the error-related negativity (ERN) (Gehring et al., 1993; Rodriguez-Fornells et al., 2002). Several evidences support the idea that these mechanisms are, at least, partially orchestrated by on-going theta oscillatory activity (4-8 Hz), as errors and response conflict consistently modulate theta power activity in the mPFC (Cavanagh et al., 2009; Cavanagh et al., 2012; Nigburg et al., 2012; Pastotter et al., 2013; van Drial et al., 2012). It has been suggested that increases in medial-fontal theta activity underlying error and response conflict represent the activation of a generic system of action-regulation that is responsible for the constant evaluation of demanding environmental situations on the outside world and online action adjustments (Cavanagh et al., 2009; Luu et al., 2003). In this sense, theta oscillatory activity may reflect a mechanism through which the mPFC, in face of adverse events, interacts with motor and other top-down control structures in order to prompt us to rapidly implement behavioral adjustments, such as cancelling or slowing down on-going inappropriate responses tendencies or implement compensatory mechanisms after actual erratic actions (Cavanagh et al., 2009; Narayanam et al, 2013; Pastorer et al., 2013).

However, the extent to which the mPFC action-monitoring system is sensitive to conflicting unattended events which cannot be consciously accessed or identified is a question that has received little attention so far. Although it has been long debated as to whether unattended visual items can be processed beyond its simple physical features, and ultimately evaluated on a goal-directed fashion, there is some consensus with regard to the idea that the influence of unattended (task-irrelevant) conflicting events on behavior is somehow unavoidable (Eriksen and Eriksen, 1974; Miller, 1991; Paquet, 2001). In the present study we examined whether unattended, non-conscious, conflicting (i.e. error-prone) events are capable of activating neural networks of action-monitoring of the mPFC and more specifically, modulate medial-frontal theta oscillatory activity.

To accomplish this goal we recorded event-related potentials (ERPs) while participants performed a novel variant of the Erikson flanker task (Eriksen & Eriksen, 1974) in which spatial attention was manipulated. Using a novel experimental design we aimed to simulate a complex visual environment in which participants selectively prioritized visual information that was more relevant to the task on-hand (*as when driving we covertly attend to the incoming traffic*) at the expense of other items that, although visible, remained oblivious (as in the given example, *the dog on the sidewalk*) but still influenced participants performance. Here, participants were instructed to covertly attend either to the left or right visual fields while bilateral flankers were parafoveally displayed at both visual fields. Participants had to respond as fast as possible to the direction of a target arrow presented on the attended side and to ignore the set of flankers exhibited at unattended parafoveal locations – which were congruent, neutral or incongruent with the attended target (see Fig. 1). The spatial arrangement of the flankers prevented participant's conscious access to the stimuli located at unattended locations (Block, 2005;Cohen et al., 2012;Lamme, 2003), such that unattended stimuli could not be explicitly reported. We predicted that even when unattended conflicting information could not be consciously identified, it would still be causally efficacious in affecting behavior and influencing the amplitude of frontocentral ERP signals as well as the power of medial-frontal theta oscillatory activity associated with action-monitoring and regulation mechanisms of the mPFC.

## METHODS

### Participants

Eighteen healthy right-handed participants from the Faculty of Psychology of the University of Barcelona participated in the experiment (7 men; age range: 18-29 years). All of them gave informed consent and were rewarded for their participation. All participants had normal to corrected vision.

## Stimuli and Design

### *Main Experiment Design*

The stimuli were presented on a 19-in. colour monitor (frequency 60 Hz) against a grey background (178 RGB) at a viewing distance of 65 cm. Bilateral stimuli, which consisted of three black arrows oriented horizontally (4.8° × 2.1 °), were displayed on the left and right side of the screen along the horizontal meridian at a distance of 5.5° from a central fixation point. The duration of the bilateral stimuli presentation was 130 ms and stimuli onset asynchrony (SOA) was set to 870 ms.

In the beginning of each block participants were instructed to attend to the stimuli presented on one side of the screen while ignoring the group of flankers presented on the unattended side. On the attended side, stimuli were formed by one central target and flankers above and below. Attended stimuli were categorized as attended Congruent, when the target and flankers pointed to the same direction, and attended Incongruent, when target and flankers had opposite directions. On The unattended side stimuli comprised 3 flankers, oriented in the same direction, that were congruent, incongruent or neutral with the direction of the target on the attended side (Fig. 1a). Participants were required to respond, as fast and accurately as possible, to the direction of the central arrow (target) presented on the attended side, by pressing one of the two buttons, using both hands, assigned to the direction of the target. It was stressed that independently of the attended side the fixation point should be maintained throughout all experimental blocks. Eye movements were monitored with an eye tracker and movements towards the attended stimuli or away from fixation were discarded off-line from the analysis (*see* methods below).

The experiment followed a 2 x 3 factorial design with factors congruency on the attended side (Congruent, C; Incongruent, I) x congruency on the unattended side (congruent, c; neutral, n; incongruent, i). The experiment included 14 blocks (50% attending to the left side and 50% to the right, counterbalanced) of 120 trials. Each condition was randomly presented equal number of times (280 trials per condition). Before the task 80 training trials were administered.

Behaviour performance was examined with ANOVAS on the mean proportion of error responses and RTs of correct responses with factors congruency attended side (Congruent, Incongruent) x congruency unattended side (congruent, neutral, incongruent).

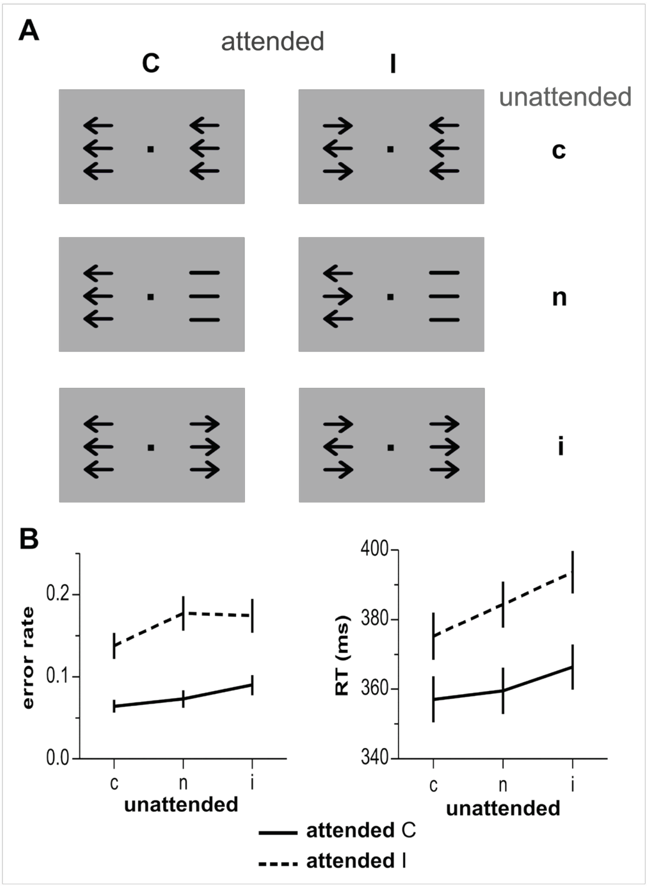


**Figure 1. (A)** Experimental task design. Example of experimental conditions in blocks in wish attention was deployed to left side of the screen. On the attended side, trials are categorized as Congruent (C) or Incongruent (I). On the unattended side stimuli can be congruent (c), neutral (n) or incongruent (i) with the target on the attended side. (**B)** Mean error rate and RTs across all conditions. Error bars represent SEM.

*Discrimination control experiment*

After the main experiment, participants performed a control discrimination task in order to evaluate the degree of attention directed to the attended stimuli and examine if the events appearing on both attended and unattended visual fields could be consciously discriminated. Two additional blocks of 156 trials (one attending to the left and the other to the right, randomly ordered) with the same stimuli and trial timing were administered. Participants were instructed to perform the control experiment exactly the same way as the main experiment. Important, participants were told that along the blocks there would be few questions to evaluate the degree of attention to the instructed attended stimuli. Specifically, they were asked to discriminate the direction of the central arrow presented on the attended or unattended side of the preceding trial. The side of the screen to which participants had to determine if left or right-pointing arrows were presented was indicated. In each block 16 questions were introduced appearing in random order - half regarding the attended stimuli and half regarding the unattended stimuli. Questions after trials in which the unattended stimuli were neutral were avoided and then counterbalanced across the different conditions. In total, 16 questions in reference to the attended and the unattended stimuli were presented.

Performance in the forced-choice discrimination task (percentage of correct responses) was accessed individually using a binominal test ($p < 0.05$). We expected that participants would not be able to discriminate unattended events above chance (0.5) since the attention to those stimuli was limited.

### Eye-movements recording and pre-processing

The participants' position of gaze was monitored using a binocular EyeLink II eye-tracking system (SR Research System, Ontario, Canada). To compensate for any head movements, we used a chin rest. Eye recordings were coordinated with the EEG recording using the Prexel (Visual Attention Lab, UMass Boston, USA) eye-tracker extension toolbox for Presentation (Neurobehavioral Systems).

The eye tracking equipment was calibrated for each participant at the beginning of each experimental block (standard 9 point calibration). The data was recorded with 500 Hz of sampling rate. Eye recordings were used to filter out those trials where eye gaze and fixation point coordinates did not match and trials in which participants did not maintain a stable fixation. A stable fixation was defined as eye movements smaller than 0.2° and slower than 0.2°/sec during the time window of 1 sec around the onset of the stimuli presentation. Trials in wish eye gaze deviates from fixation were discarded from the behavioural and EEG analysis. In the main experiment 13 ± 7 % (mean + SD) of all the trials were rejected offline. All participants had a rejection rate smaller than 25 %. For the control task, two participants were excluded from the analysis due to an excessive number of saccades away from the fixation point (>30 % of the trials). For the remaining 16 participants, on average 13 questions regarding the attended and unattended stimuli were analyzed (max:14; min:12).

### EEG Recording and ERP analysis

EEG was recorded from tin electrodes mounted in an elastic cap and located at 28 standard positions (Fz, Fcz, F7/8, F3/4, Fc1/2 Fc5/6, Cz, C3/4, T3/T4, T5/T6 CP1/2, CP5/6, Pz, P3/4, PO1/2, O1/2). All scalp electrodes were referenced offline to the mean activity of the left mastoid. EOG was recorded with electrodes located bellow the right eye (vertical EOG) and electrodes placed on the outer canthus of each eye (horizontal EOG). Both vertical and horizontal EOG were used for artefact rejection.

Impedances were kept below 5 kOhm. The EEG signals were filtered with band-pass of 0.1–70 Hz (half-amplitude cutoffs) and digitized at a rate of 250 Hz. EEG was low-pass filtered offline at $<$ 14 Hz for ERP analysis.

Only those trials in which the fixation was maintained (trials pre-selected from the eye movement analysis) were analysed. ERP effects associated with response conflict and response errors (ERN) were studied by extracting epochs from -1000 ms to 1000 ms locked to the onset of the response (baseline period was defined from -400 to -200 ms prior to the response onset). ERP responses associated with spatial attention mechanisms were also inspected by extracting epochs of 1000 ms locked to the onset of the stimulus presentation (baseline corrected from -100 ms to the onset of stimulus presentation).

For both stimulus-locked and response-locked analysis, trials exceeding ± 75-100 µV in both EEG and vertical EOG during the epoch window were rejected offline (thresholds were adjusted individually for each participant). Trials with horizontal eye movements that exceeded 2 standard deviations relative to baseline activity were also removed. For each participant we centered our analysis on trials with correct responses. Additionally, trials following resting periods, error responses, attended Incongruent trials (i.e. high response conflict trials) and trials with RTs shorter than 200 ms or longer than 2.5 standard deviation of the individual RT mean, were excluded from the analysis.

### *ERP analysis on spatial attention effects*

In order to ensure that participants consistently maintained their attention to the instructed target stimuli on the attended side we explored effects of spatial attention on visual evoked ERPs (Hillyard & Anllo-Vento, 1998; Mangun, Hillyard, & Luck, 1993).

Trials in which attention was directed to the left side were separated from trials in which attention was directed to the right side. No difference between the number of trials analysed attended left (M + SEM: 501 ± 18) and attended right (M + SEM: 505 ± 22) was seen ($t(17) < 1$). Covert deployment of visual attention to the target stimuli on the attended side was associated with an enhancement of early visual ERPs on the time windows of 70-110 ms (P1) and 220-260 ms (N2pc) on posterior sites of the cortex contralateral to the attended side (Fig. 2). For statistical analysis we selected two ROIs of three sensors over posterior regions of the scalp: on the right hemisphere (RH: P4, P02, T6) and on the left hemisphere (LH: P3, P01, T5) (the selected ROIs are highlighted on the topographical map of the Fig. 2b, left panel). Statistical effects were tested by ANOVAs with factors attended side (Right, Left) x hemisphere recording site (ipsilateral, contralateral) on the mean amplitude of the P1 and N2pc components.

### *ERP analysis related to response conflict and error-monitoring*

To examine whether unattended events would modulate electrophysiological markers of response conflict we conducted our analysis on correct responses instead of looking at overt errors. It was observed that trials with increased conflict on both attended and unattended side (i.e. incongruent trials) elicited a negative ERP component with frontocentral distribution arising 130-80 ms before the response onset.

EEG data were additionally reanalysed using a band pass filter [3-9 Hz] in order to retain theta-band specific ERP activity. Accordingly, it has been demonstrated that frontocentral negative ERPs, associated with error/conflict monitoring, reflect neural oscillations in the theta band (Cavanagh et al., 2012). Furthermore, this method is especially advantageous because it allowed us to remove the positive slow wave deflection in which the frontocentral negative component is superimposed (for a similar approach sees Luu et al., 2001).Statistical analysis was conducted on the mean amplitude of this negative ERP with ANOVAs with factors congruency attended side (Congruent, Incongruent) x congruency unattended side (congruent, neutral, incongruent) x electrode location (frontocentral sites: Fz, Fcz).

Additionally, we examined the amplitude of the ERN component after overt errors on unattended congruent, incongruent and neutral trials in order to compare the timing between the ERPs related to response conflict and error commission during unattended conflicting trials. Moreover this analysis was crucial to examine if both the ERN and the frontocentral negative ERP stem from overlapping generator structures in the mPFC (see electrical source analysis below) (van Veen & Carter, 2002; Yeung et al, 2004;). The ERN was calculated as the difference between error and correct responses within a window of 40-80 ms. The mean amplitude of the difference waveform of the ERN was tested with ANOVAs with factors congruency unattended side (congruent, neutral, incongruent) x electrode location (frontocentral sites: Fz, Fcz).

### **Source Analysis of conflict and error-related ERP activity**

Brain Electric Source Analysis (BESA 2000 version 5.3) (Scherg, 1990) was additionally used to determine the source of the ERP signal observed during conflicting attended and unattended events. We predicted that the frontocentral negative component in both attended and unattended conflicting events would overlap with the ERN dipole (overt error trials) in the mPFC, suggesting the involvement of this area on error and conflict monitoring to attended sources of conflict and, importantly, to unattended conflicting events too.

BESA algorithm computes the location and the orientation of multiple equivalent dipolar sources by calculating the voltage scalp distribution that would be produced for a given dipole model (forward solution) and comparing it with the original scalp distribution. Source localization analysis was carried out on the ERP difference waveforms obtained subtracting (i) error minus correct (ERN response) (ii) attended Incongruent minus attended Congruent and (iii) unattended incongruent minus unattended congruent trials, after applying a band-pass filter in the theta range. First, we fitted one single dipole in the ACC/pre-SMA areas, which had shown significant error and conflict-related fMRI activations in previous studies (for reviews see Ridderinkhof et al., 2004;Ullsperger et al., 2014), within the ERN component (0-100 ms). We subsequently tested the strength of this dipole model estimating the neural sources of the ERP components in attended Incongruent and unattended incongruent conditions. These intervals were defined within a ±30 ms time window centered on the peak-latency of EEG activity. The latencies of major peaks in the dipole source waveform were also taken as indices of neural response timing, which were concomitant with the peak of ERP activity. The final locations of each dipole in the group-average BESA model were projected on mean structural T1 MRI image of 24 individuals and converted into Talairach coordinates in the standard Montreal brain (Talairach and Tournoux, 1988).

### Time-frequency analysis related to response conflict

Furthermore time-frequency analysis was performed on the EEG activity for each trial in 4-sec epochs (from -2 sec to 2 sec locked to the onset of the response) using seven-cycle complex Morlet wavelets (Marco-Pallares et al., 2008). In order to investigate trial-by-trial modulations of power in the theta band (4-8 Hz), changes in time-varying energy (square of the convolution between wavelet and signal) in the studied frequencies (1 to 40 Hz, linear increase) with respect to baseline were computed for each trial and average for each participant before performing a grand average. The mean increase/decrease in theta power for each condition was inspected during the temporal window of -200 to 0 ms (locked to the response onset) using ANOVAs with factors congruency of attended side (Congruent, Incongruent) x congruency unattended side (congruent, neutral, incongruent) at the Fcz electrode, which was the electrode with maximal theta power activity in trials with enhanced response conflict.

## RESULTS

### Behavioural Results

*Discrimination task*

The results from the forced-choice discrimination task showed that the mean percentage of correct responses for stimuli presented on the attended side was 93 ± 8%; and for unattended stimuli 61 ± 11% ($t(15) = 9.2$, $p < 0.001$). In all cases, participants were unable to discriminate the stimuli on the unattended side above the chance level (0.5) (binomial test, $p < 0.05$).

*Main Experiment*

Consistent with previous behavioural results in the flanker task, it was observed that the mean RT on attended Incongruent trials (M ± SD: 384 ± 27 ms) was higher than attended Congruent trials (361 ± 28 ms) (main effect of congruency attended side: $F(1,17) = 203.3$, $p < 0.0001$) (Fig 1b, right panel). Notably, participants' RTs were also affected by the congruency of unattended flankers (main effect of congruency unattended side: $F(1,17) = 56.6$, $p < 0.001$). As observed in Fig. 1b unattended incongruent conditions caused greater response conflict, reflected by slower RTs (380 ± 26 ms), compared to unattended neutral (371 ± 28 ms) and unattended congruent conditions (366 ± 28 ms). Further paired t-tests between the three conditions (unattended c, n and i) confirmed the statistical differences between unattended incongruent and neutral conditions ($t(17) = 7.2$, $p < 0.001$); unattended incongruent and congruent conditions ($t(17) = 9.5$, $p < 0.001$); and unattended congruent and neutral conditions ($t(17) = 4.3$, $p < 0.001$).

Furthermore, analysis on the mean error rate showed that participants were less accurate on attended Incongruent trials (M ± SD: 0.16 ± 0.07) compared to attended Congruent ones (0.08 ± 0.04) (main effect of congruency attended side: $F(1,17) = 55$, $p < 0.0001$). Also, it was observed a modulation of the error rate as function of the congruency on the unattended side ($F(2,34) = 9.4$, $p < 0.001$) (Fig. 1b, left panel), which shows that unattended conflicting events are more prone to errors than non-conflict or neutral unattended events. Subsequent paired *t-tests* showed that unattended incongruent trials generated more errors (M ± SD: 0.13 ± 0.07) than unattended congruent trials (0.10 ± 0.05) ($t(17) = 3.8$, $p < 0.001$) and unattended congruent flankers caused less errors than neutral flankers (0.12 ± 0.06) ($t(17) = 3$, $p < 0.01$). No differences were found on the comparison unattended incongruent - neutral conditions ($t(17) = 1.1$, $p > 0.05$) (Fig. 1b, left panel).

In sum, the behaviour results showed that not only attended sources of conflict, but also unattended (unconscious) conflictive events, elicited response conflict reflected by slower RTs and higher error rates.

## ERP Results

### *Spatial attention effects*

It was observed that covert deployment of attention to the attended stimuli (vs. unattended stimuli) was associated with increased activity of the P1 and N2pc components on posterior sites of the cortex contralateral to the attended side (Fig. 2).

ANOVAS with factors attended side x hemisphere recording site (ipsilateral, contralateral) on the mean amplitude of the P1(70-110 ms) and N2pc (220-260 ms) confirmed the contralateral lateralization effect (main effect of hemisphere recording site) for the P1 ($F(1,17) = 69.4$, $p < 0.0001$) and the N2pc ($F(17) = 25.9$, $p <$ 0001) components (Fig. 2a). On topographical maps of Figure 2b is clear the enhancement of both components on the parietal-occipital recording sites contralateral to the attended side. For both components, neither a main effect regarding the attended side nor an interaction attended side x hemisphere recording site were observed (both $Fs<1$). These results are important because they showed that consistently participants maintained their attention to the instructed target stimuli.

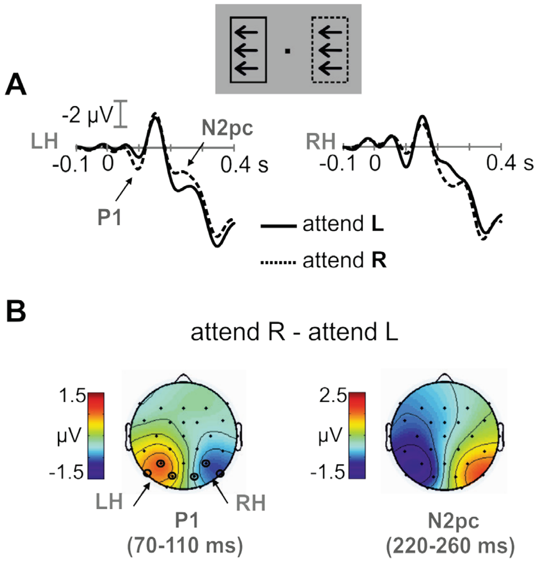


**Figure 2.** **(A)** ERPS regarding spatial attention effects on early visual ERPs. Spatial attention to one side of the screen was associated with the increase of the P1 and N2pc components on the contralateral posterior sites of the scalp. (**B)** Scalp distribution for the spatial attention effect (attend Right - attend Left). Black disks on the topographical maps denote the ROI of parietal-occipital recording electrodes on the right (RH) and left hemisphere (LH).

### *Frontocentral ERP activations in in response to unattended conflicting events and overt error responses*

Our results showed that conflicting stimuli on both attended and unattended locations, i.e incongruent flankers, triggered a frontocentral negative component, peaking at about 100 ms before the onset of the correct responses, that reflects on-going activity on the theta band (Fig. 3 and 4).

The results from the ANOVAS confirmed an enhanced frontal negative activity in attended Incongruent correct trials compared with Congruent correct ones at the Fcz electrode, as demonstrated by the interaction congruency on the attended side x electrode location ($F(1,17) = 4.8$, $p < 0.05$). This difference between attended Incongruent and attended Congruent trials at the Fcz electrode was confirmed by pairwise comparisons ($t(17) = 4.9$, $p < 0.0001$; attended Incongruent (M ± SEM), 0.9 µV ± 0.05; attended Congruent, 1.8 ± 0.05 µV; note that absolute values are positive because the negative component is superimposed on a slow positive deflection) (Fig. 3a). The frontal negative component was also shown in response to conflicting stimuli on the unattended side (main effect of congruency on the unattended side: $F(1,17) = 3.6$, $p < 0.05$), being greater during unattended incongruent trials (Fig. 3b; *see* also the topographical maps illustrating in Fig. 4a). Statistical differences regarding the amplitude of this ERP at Fcz (average across attended Congruent and attended Incongruent conditions) in trials with different degrees of conflict on the unattended side (congruent, neutral and incongruent) were tested with paired *t-test* comparisons. The amplitude of this negative ERP at Fcz (which was the electrode with maximal attended conflict-related effects) was greater in unattended incongruent than congruent conditions ($t(17) = 2.5$, $p < 0.05$) and, although in a less degree, than neutral conditions ($t(17) = 2.1$, $p <= 0.05$) (Fig. 4c, left panel). No differences between unattended congruent and neutral conditions were found ($t(17) < 1$).

After filtering the data with a band-pass filter in the theta range we observed, again, the same negative frontocentral response, around 100 ms before response onset, in trials with higher conflict (Fig. 3b). A main effect of congruency on the attended side ($F(1,17) = 6.9$, $p < 0.05$) confirmed an increased negative cortical activity in attended Incongruent trials compared to attended Congruent trials (M ± SEM: attended I, -1.4 µV ± 0.03; attended C, -0.9 ± 0.02 µV). Again, the congruency on the unattended side affected the amplitude of this frontocentral ERP signal ($F(2,34) = 5.2$, $p < 0.05$). Further pairwise comparisons confirmed the enhanced negativity in unattended incongruent trials compared to congruent *(*$t(17) = -3$, $p < 0.01$) and neutral conditions (marginal increase: $t(17) = 2$, $p < 0.06$). No significant differences between unattended congruent and neutral conditions were found ($t(17) = 1.5$, $p > 0.05$) (Fig. 4c, right panel).

In addition, we inspected ERP responses following error trials in order to compare the temporal dynamics of response conflict monitoring on correct responses and error-monitoring during overt errors (see Fig. 5). The ERN after overt error trials showed a frontocentral distribution being maximal at the Fcz electrode (main effect of electrode: $F(1,17) = 4.7$, $p < 0.05$) (Fig. 5a). However, in contrast with the frontocentral negative ERP appearing before the onset of correct responses in conflicting trials, the ERN was not modulated by the congruency on the unattended side (bottom Fig. 5a). Both the congruency unattended side or the interaction congruency unattended x electrode location were not significant (both $Fs < 1$). Then, the data was reanalysed with a band pass filter in the theta range (Fig. 5b). Again the ERN was maximal at the Fcz electrode (main effect of electrode: $F(1,17) = 19.4$, $p < 0.0001$) and not modulated by the congruency on the unattended side, the main effect of congruency unattended side and the interaction congruency unattended x electrode location were not significant (both $Fs < 1$).

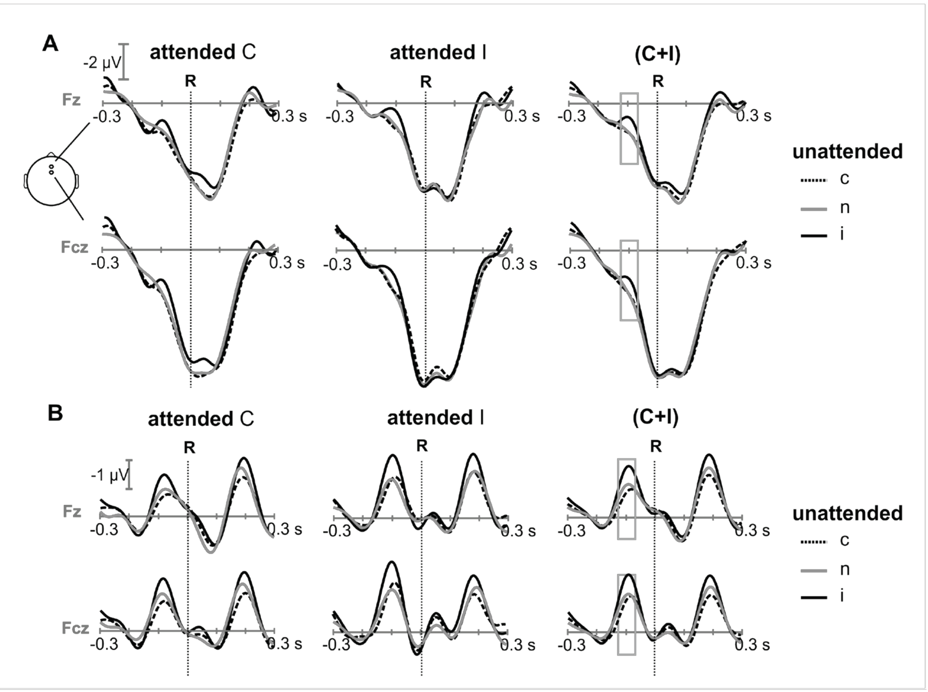


**Figure 3.** R-Locked ERPs, over frontocentral electrodes showing pre-response conflict during correct responses. Conflict effects induced by unattended stimuli – congruent (dark dashed), neutral (grey solid) and incongruent (dark solid) – are plotted separately for attended Congruent (C, left) and attended Incongruent (I, middle) conditions. On the right panel is plotted the average (attended C+I) to highlight the congruency effects resulting from unattended events. **(A)** Grand-average waveforms low-pass filtered at < 14 Hz. **(B)** Grand-average ERPs band-pass filtered in the theta range (3-9 Hz). Note a negative component around 100 ms before response onset (R, marked by vertical dot lines) that is more pronounced in unattended incongruent events (grey box).

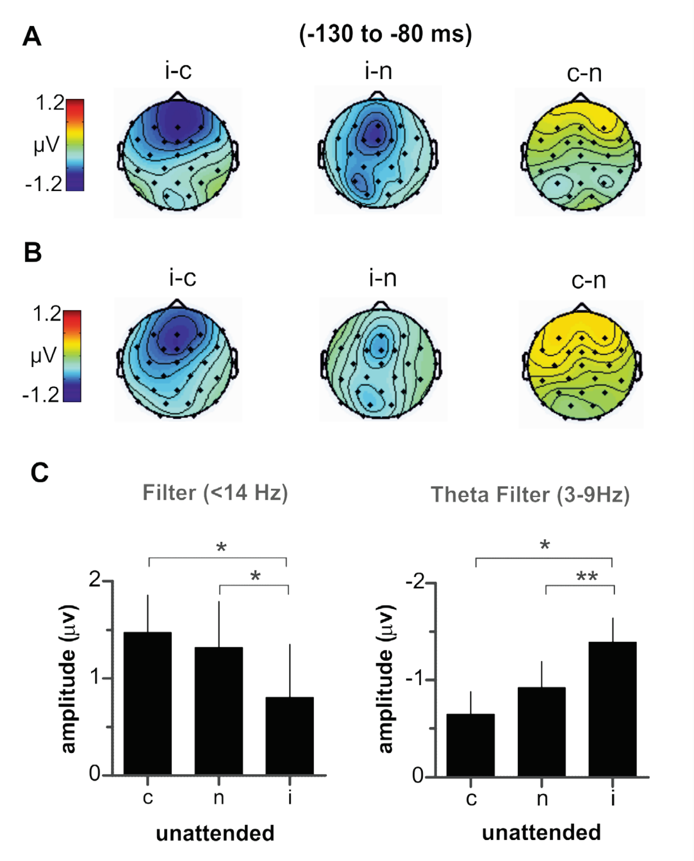


**Figure 4.** Frontocentral ERP activations during unattended conflicting trials. **(A)** Scalp topography (ERPs low-pass filtered at <14 Hz) of the conflict effects induced by unattended events (from left to right, paired *t-test*: unattended incongruent-congruent; unattended incongruent-neutral; unattended congruent- neutral). **(B)** Scalp topography (ERPs band-pass filtered at theta range) of the conflict effects induced by unattended events (paired *t-test*). **(C)** *T-test* comparisons on the mean amplitude of the negative frontocentral ERP between the unattended conditions. * $p<0.05$; ** $p<0.06$; Error bars represent SEM.

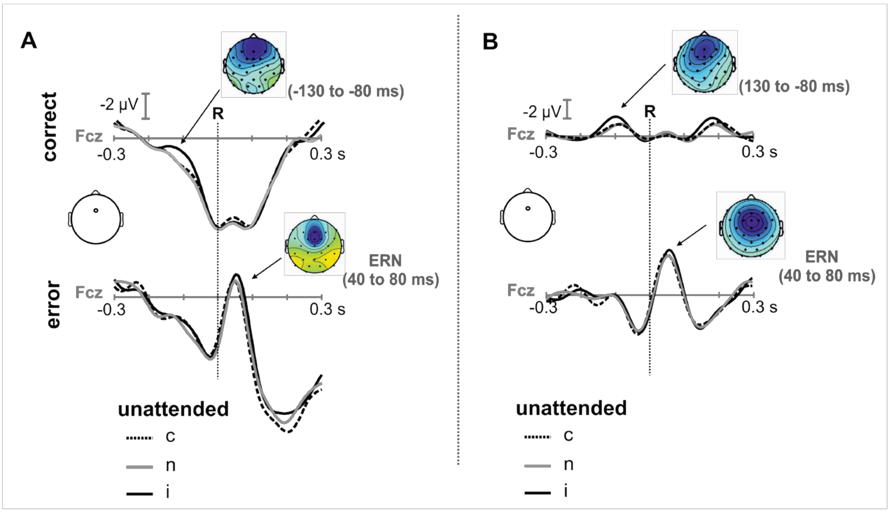


**Figure 5.** R-Locked ERPs for correct (top) and error (bottom) responses as function of the congruency of unattended events. Note the different timing of the negative frontocentral ERP (top) associated with pre-response conflict and the ERP related to post-error commission (the ERN on the bottom of the figure). (**A**) Grand-average waveforms for the ERPs using a low-pass filtered at <14 Hz. **(B)** Grand-average waveforms for the ERPs with a band-pass filter in the theta range. Also are depicted the topographical maps for the significant effect of unattended congruency (incongruent-congruent) on pre-response dynamics and effects of error commission (error-correct trials). Note that in contrast with the pre-response negative component the ERN was not modulated by unattended congruency.

### Source localization of the frontocentral ERPs related to response conflict and overt errors

Source localization results showed that a single source in the right ACC (Talairach coordinates (*x, y, z*): 3.4, 8.2, 45.8) accounted for 88.8% of the variance of the ERN response in error trials (temporal window of 0-100 ms post-response), with a residual variance (RV) of 11.2%. When the time window was restricted to a ±20 ms interval with respect to the ERN peak (40-80 ms), RV decreased until 4.9% (Fig. 6A). Additionally, this single dipole model was found to explain 87.7% of the variance of the negative frontocentral ERP component associated with response conflict in attended Incongruent trials (time window of peak activity: -160 to -100 ms) and for the 88.2% of the variance in unattended incongruent trials (time window of peak activity: -130 to -70 ms) (Fig. 6). Source waveforms were additionally obtained for each dipole between -200 and 400 ms post-response (right panel Fig. 6). Notably, the source waveforms showed similar pre-response peak-latency in both attended and unattended conflicting conditions. Our results confirm our predictions by showing that the frontocentral negative responses in the theta band in both attended and unattended conflicting events overlap within the same neural networks involved in conflict and error monitoring in the mPFC.

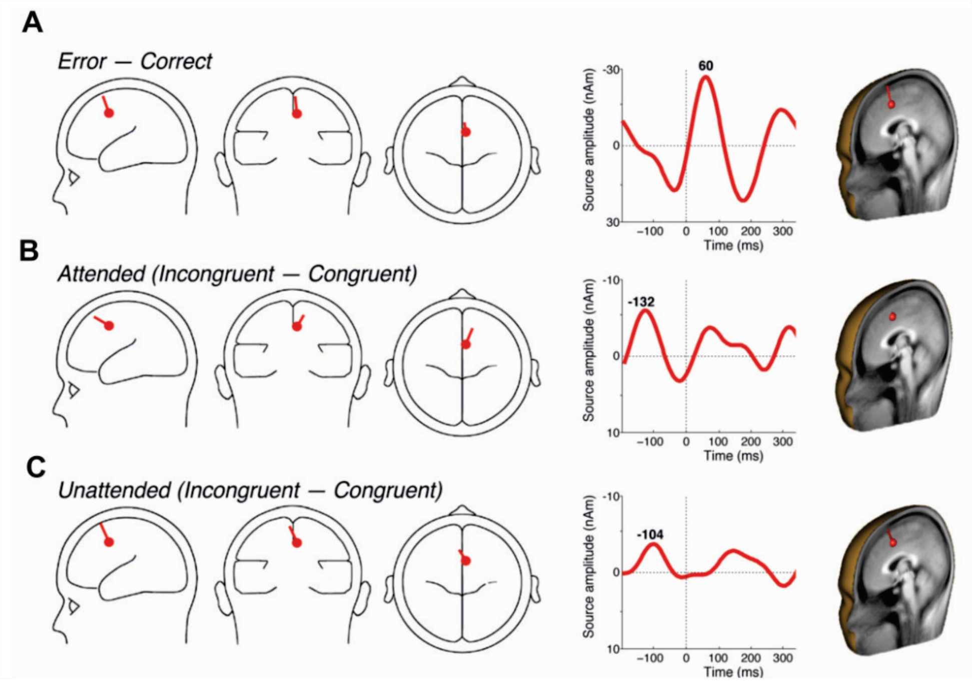


**Figure 6.** Source localization (BESA) analysis revealing a common neural source in the ACC for the monitoring of errors and conflict during attended and unattended incongruent conditions. **(A)** single dipole in the right ACC (red; $x = 3.4$, $y = 8.2$, $z = 45.8$) was found for the ERN response during error trials. The same dipole explained 87.7% of the variance of the negative frontocentral ERP during conflicting attended trials **(B)** and 88.2% of the variance during conflicting unattended trials **(C)**. Time-course of each computed dipole is represented in the source waveform on the right side of the figure.

### Medial-frontal theta power increases in unattended conflicting events

To investigate spectral dynamics of response conflict, time-frequency spectral power was examined during attended(Incongruent vs Congruent) and unattended (congruent, neutral, incongruent) conditions at Fcz electrode.

We observed greater medial-frontal theta power (4-8 Hz) during the 200 ms prior to the response onset in attended Incongruent trials compared with attended Congruent ($F(1,17) = 7.4$, $p < 0.05$) (Fig. 7a). A main effect of congruency on the unattended side was also encountered ($F(2,34) = 4.2$, $p < 0.05$) (Fig. 7b and 7c). The interaction congruency attended side x congruency unattended side was not significant ($F(2,34) < 1$). In agreement with the ERP results, medial-frontal theta power was enhanced in unattended incongruent than unattended congruent conditions ($t(17) = 4$, $p < 0.005$) (Fig. 7c). However, no differences between unattended incongruent and unattended neutral conditions ($t(17) = 1.2$, $p > 0.05$) or unattended neutral and unattended congruent conditions were found ($t(17) = 1.3$, $p > 0.05$).

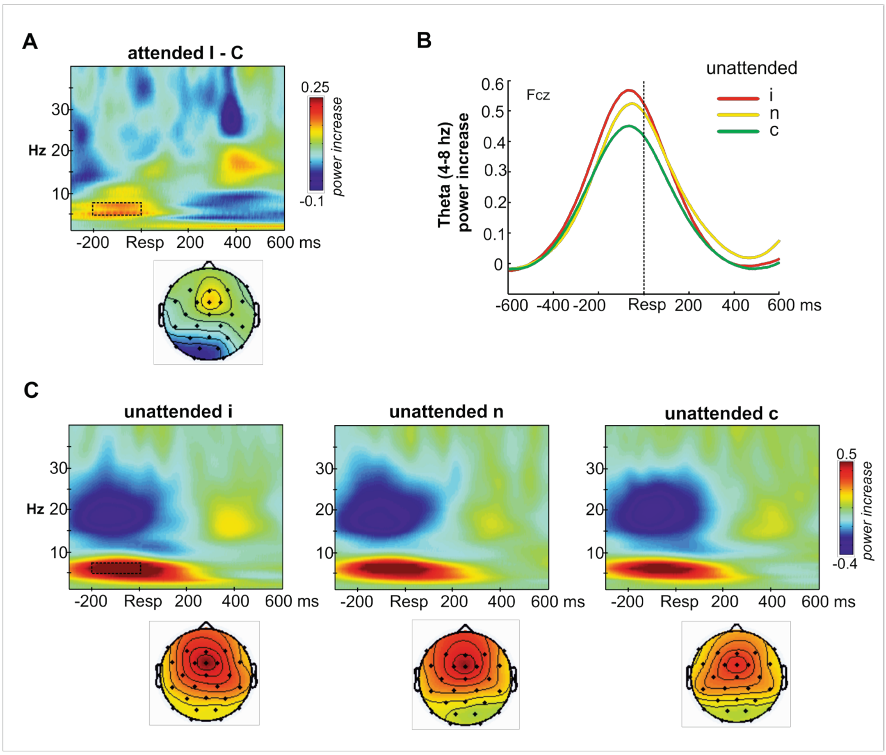


**Figure 7.** (**A**) Media-frontal theta power increase at Fcz electrode prior to response onset in attended incongruent conditions (attended Incongruent-Congruent) and associated topographical map. (**B**) Line plots of Fcz theta power activity over time for the different unattended conditions (i, incongruent; n, neutral; c, congruent). (**C**) Medial-frontal theta power increase at Fcz electrode for the different unattended conditions and respective topographical maps.

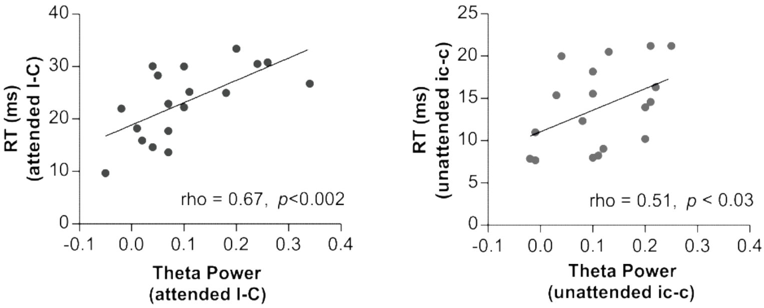


**Figure 8.** Spearmen correlations calculated between the medial-frontal theta power at Fcz electrode during attended (left) unattended (right) incongruent trials (incongruent-congruent) and the RT slowing.

We next examined whether medial-frontal theta power increases during attended and unattended conflicting trials predicted participants' slowing down. Spearmen correlations were computed between the mean theta power at Fcz during attended (difference Incongruent vs. Congruent) and unattended (incongruent vs. congruent) conflicting conditions and the RT in attended and unattended incongruent trials (difference attended I-C and unattended i-c, respectively) (Fig. 8). It was found that enhanced medial-frontal theta power in attended Incongruent trials significantly correlated with the RT slowing ($r = 0.67$, $p = 0.0022$). Remarkably, medial-frontal theta power increases in unattended incongruent conditions also correlated with the RT slowing down during unattended conflict trials (unattended i-c: $r = 0.51$, $p = 0.03$).

## DISCUSSION

Earlier theories have stressed the role of the mPFC during voluntary and attentive control of action, and especially in monitoring and resolving situations where inappropriate or error-prone response tendencies are triggered (Jack and Shallice, 2001; Posner et al., 1994; Posner & DiGirolamo, 1998). To the best of our knowledge, no previous study has examined whether the mPFC also participates in the monitoring and regulation of events that remain unattended and which neural oscillatory mechanisms may sustain these processes.

In the present study, using combined ERP, time-frequency and electrical source (BESA) analysis, it was showed that conflictive stimuli presented at parafoveal unattended locations, which could not be consciously reported, still influenced participants' behaviour (by means of increased RTs and error rates) and modulated mPFC theta oscillatory activity. In both attended and unattended incongruent conditions, a negative frontocentral ERP, reflecting on-going theta oscillatory activity, and an increase of medial-frontal theta power right before the onset of correct, yet conflicting, responses were observed. Important, electrical source analyses localized the source of this theta-band specific frontocentral ERP in the mPFC. A positive correlation between the conflict-induced medial-frontal theta power and the magnitude of participant's slowdown, further suggested that theta oscillations in the mPFC not only participate in the monitoring of conflicting events presented on the environment, but also regulate inappropriate response tendencies triggered by external events that are processed without much attentional control.

In the present study we set up a novel experimental design in order to examine response conflict processes triggered by visual events that were paravofeally processed under conditions of limited visual attention. It is important to mention that subjects were not deceived from a conscious phenomenal visual experience (see Lamme et al., 2003) of the events located at both attended and unattended parafoveal locations; yet, they showed drastic problems to give accurate reports about the nature of stimuli presented at the unattended locations. Participants performance during the force-choice discrimination task was at chance level, which shows that the monitoring of the unattended events unfold, at least, partially, without an explicit knowledge or a metacognitive representation about the sources of conflict. These findings follow theoretical frameworks that propose that visual attention is an important requisite to access conscious representations of perceived stimuli on the outside world (Lamme, 2003; Posner, 1994), influencing the likelihood of accurate reports about those stimuli (Lamme, 2003; Wyart and Tallon-Baudry, 2008; Cohen et al., 2012).

Neurophysiological support for this idea came from the observation that covert deployment of attention towards the target on the attended locations was associated with an increase of the P1 and N2pc components on contralateral posterior sites of the cortex (Fig 2). As both ERP signals have been interpreted as a neural index of visual attention (Hillyard and Anllo-Vento, 1998; Mangun et al., 1993), this control analysis showed that participants consistently maintained their attention to the instructed target stimuli on the attended side. Therefore, selective attention to the attended target stimuli may have biased perceptual judgements of the unattended stimuli by selectively granting priority in processing of information of attended (task-relevant) stimuli (Hawkins et al., 1990; Posner et al., 1994; Hillyard and Anllo-Vento, 1998).

Although stimuli presented at unattended locations may have generate weaker representations and therefore remained unavailable for conscious report, they still influenced participant’s behavior. Both attended and unattended incongruent trials were related with slower RTs and higher error rates (Fig 1b). Thus the present results are in agreement with previous studies showing the difficulty of abolishing completely the flanker compatibility effect event when conflicting information is presented in unattended and distant locations, which was the case in the present study (Eriksen and Eriksen, 1974; Lavie, 1995; Miller, 1991; Paquet, 2001).

Notably, unattended incongruent stimuli compared to congruent and neutral unattended stimuli were associated with a negative component in the pre-response phase (~ 100 ms prior the response onset) that, by timing and scalp topography, is consistent with a frontal N2 component – an ERP correlate of a conflict monitoring in the mPFC (van Veen & Carter, 2002; Yeung et al., 2004). By applying a band-pass filter (4-9 Hz) in the EEG data, we were able to confirm that this frontocentral negative ERP during both attended and unattended conflicting conditions reflects ongoing medial-frontal oscillatory activity in the theta band. Source localization analysis showed that this frontocentral theta-specific ERP originated in the mPFC, a key structure in error and conflict monitoring and action selection (Botvinick, 2001; Carter, 1998; Holroyd et al., 2004; Ridderinkhoff, 2004).

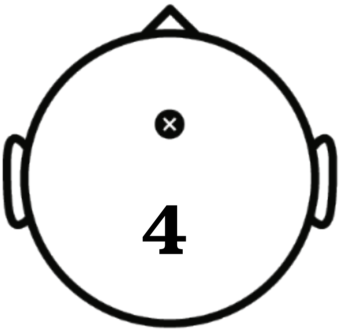
The observation of enhanced medial-frontal theta power during response conflict replicates earlier electrophysiological findings (Cavanagh et al., 2012; Nigburg et al., 2012; Pastorer et al., 2013). However, to the best of our knowledge, no previous study has reported medial-frontal theta activity in the mPFC in response to unattended and unconscious conflicting stimuli.

It has been proposed that theta oscillations support conflict/error monitoring and consequent compensatory mechanisms by coordinating the mPFC with other control and motor networks supporting flexible and adaptive decision process (Cavanagh et al., 2009; Cavanagh et al., 2012; Luu et al., 2004; Narayanan et al., 2013; Womelsdorf et al., 2010). For instance, a recent study in humans and rodents has demonstrated that post-error compensatory mechanisms strongly rely on the coordination of slow theta oscillations between the mPFC and the motor cortex (Narayanan et al., 2013), extending previous neuroimaging studies showing that error-related mPFC activity predicts post-error slowing mechanisms as well as consequent decreases in the motor activity (Danielmeier et al., 2011; King et al., 2010).

In agreement with these findings here we showed that the magnitude of media-frontal theta power increase on both attended and unattended conflicting trials was correlated with the degree participants slowing down. Given the present results, it is very likely that medial-frontal theta activity arising prior to action initiation is involved in the cancelation or slowing down of prepotent erroneous actions even when those actions were activated automatically by stimuli that remained outside the scope of attentional control. As an alternative, but not incompatible, explanation, it is possible that the observed enhancement of theta activity was associated with a larger number of partial errors during conflicting unattended trials (Cohen & van Gaal, 2014). Partial errors make reference to subthreshold muscle twitches during correct responses that are ensued by suppressing on-going erroneous response tendencies (Gratton et al., 1988; Allain et al., 2009). Although this hypothesis is hard to disentangle since electromyography (EMG) activity was not recorded, our results suggest that somewhat theta activity in the mPFC is involved in some sort of regulatory processes in order to supress inappropriate response tendencies activated by conflicting environmental events.  Our results lend support to the idea that such regulatory processes are instantiated in prefrontal networks irrespective to the degree of attentional control (Jackson & Shallice, 2001; Posner 1994; Posner & DiGirolamo, 1998).

Finally, our data is in agreement with recent observations that the mPFC participates in the monitoring of erroneous actions (Hester et al., 2005; Nieuwenhuis, et al., 2001) and conflicting events (D'Ostilio & Garreux, 2011; Sumner et al., 2007; van Gaal et al., 2008; 2011) which cannot be consciously accessed. Our study, however, is different from previous studies that have used subliminal priming paradigms to study unconscious processes of action-monitoring (D'Ostillio and Garraux, 2012;van Gaal et al., 2010). Here, even when participants might have remained unaware of the specific details of the information presented at the unattended side, stimuli were still visible, which differs from subliminal presentations. We believe that the paradigm presented here is more similar to real-life scenarios in which our senses are constantly bombarded by visible information out of the focus of our attention that can influence our behavior and, thereby, activate neural networks related to action monitoring and regulation in a way similar to an exogenous capture of the action-monitoring system by unattended stimuli.

In sum, the present findings demonstrate that mPFC activity related to response conflict monitoring, classically associated with conscious and attentional control processes, can take place in response to unattended conflicting events, irrespective of the degree of conscious representation of the sources of conflict. The present study provides electrophysiological evidences that oscillatory theta activity in the mPFC is a plausible neural mechanism though which flexible monitoring and compensation are exerted to suppress and regulate potentially inappropriate actions that are automatically triggered by environmental stimuli to which we are oblivious.



# 4

# Study 2

## ERP evidence of adaptive changes in error processing and attentional control during rhythm learning [*]

---

[*] This study corresponds to:
Padrão, G., Penhune, V., de Diego-Balaguer, R., Marco-Pallarés, J., Rodríguez-Fornells, A. (2014) ERP evidence of adaptive changes in error processing and attentional control during rhythm synchronization learning. NeuroImage, (**in Press**)

## INTRODUCTION

For most people synchronizing movement with a musical rhythm by dancing, clapping or tapping their feet is natural and requires no special effort. However, expert musical performance requires the production of complex rhythms that are mastered through practice. When learning a new piece, the musician first has to form an internal template of the temporal events of the rhythm to be able to plan and execute the correct sequence of movements (Pfordresher & Palmer, 2006; Pfordresher, Palmer, & Jungers, 2007). Furthermore, these internal representations of the rhythm might be used to monitor the motor output and to track errors in order to improve performance. Early in learning, this template may be less well determined but with practice and feedback, it becomes more precise.

Error processing in humans is thought to be mediated by a system operating in the medial-frontal cortex (MFC) (Botvinick et al., 2001; Holroyd & Coles, 2002; Ridderinkhof et al., 2004). The contributions of this system during music performance have only recently begun to be explored. Behavioral and ERP studies have shown that musicians are able to plan several notes in advance (Pfordresher & Palmer, 2006; Pfordresher et al., 2007) and that they can detect upcoming errors even before the action is initiated or auditory feedback available (Maidhof, Rieger, Prinz, & Koelsch, 2010; Ruiz et al., 2009). Thus it has been suggested that during skilled performance, error-monitoring does not rely on external feedback but is mainly guided by internal, feed-forward models of the motor plan (Desmurguet & Gratfton, 2000; Wolpert et al., 1995), which enables fast error-correction processes (Rabbit, 1966; Rodriguez-Fornells et al., 2002

These studies have provided suggestive evidence about the possible engagement of error-monitoring mechanisms during skilled music performance, but they do not address how error processes might evolve during the acquisition of a new skill. To address this question we examined the process of learning musical rhythmic patterns in untrained individuals and measured changes in ERP signals associated with error-monitoring and error-awareness.

Studying error processing in the context of rhythm learning is particularly interesting for two reasons. First, while error-monitoring has been hypothesized to be related to skill acquisition (Adams, 1971; Palmer & Drake, 1997), to the best of our knowledge no previous ERP studies have examined how error-monitoring and error-awareness are influenced by learning during the acquisition of either musical and general motor skills. Second, the type of errors produced during skill learning differs from the type of errors studied in standard reaction time tasks (Eriksen & Eriksen, 1974) in which errors are categorical. During skill learning errors are often relative: the right response at the wrong time, or awkwardly performed. In addition, the perception of error commission, error-awareness, may depend on the accuracy of the internal representation of the action to be executed, and this representation changes with learning. Therefore we sought to examine possible changes in the error-monitoring and error-awareness systems with learning in a context that mirrors many situations in which we acquired new motor skills.

EEG studies examining performance on reaction time tasks have consistently shown that erroneous responses lead to an early negative component in frontocentral electrodes appearing immediately after error commission (within 0-100 *ms* after error onset), the error-related negativity (ERN) (Falkeinstein et al., 1990; Gehring et al., 1993). Evidence suggests that the ERN is generated in the MFC, specifically in the anterior cingulate cortex (ACC) and pre-SMA (Debener et al., 2005; Dehaene et al., 1994; Yeung et al., 2004; for a review see Ullsperguer et al., 2014), and reflects brain mechanisms sensitive to inappropriate action-sets and conflicting information. It has been proposed that the ERN may represent a neural signature of the implementation of top-down control mechanisms, such as the reallocation of attention or the adjustment of motor thresholds to promote post-error adaptions (Botvinick et al., 2001; Kerns et al., 2004; Marco-Pallares et al. 2008; Ridderinkhof et al., 2004). While the ERN may reflect a neural signal for fast and automatic processes of error-monitoring irrespective to the degree of error-awareness (Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001), conscious error perception has been associated to a slow and sustained centroparietal positive ERP response that occurs 300 to 500 ms after the error onset, the error-related positivity (Pe) (Murphy, Robertson, Allen, Hester, & O'Connell, 2012; Nieuwenhuis et al., 2001; O'Connell et al., 2007; Overbeek et al., 2005; Steinhauser & Yeung, 2010). Important, using principal components analysis (Arbel & Donchin, 2009) it has been shown that the Pe is composed of two different subcomponents: an early frontocentral component and a later centroparietal component which in terms of latency and scalp distribution resembles the P3b as a neural response to salient target stimuli in oddball tasks (Overbeek et al., 2005; Ridderinkhof, Ramautar, & Wijnen, 2009). Evidence from signal detection theory (Hillyard, Squires, Bauer, & Lindsay, 1971; Woods, Hillyard, Courchesne, & Galambos, 1980) and, more recently, from the context of error detection (Murphy et al., 2012; Steinhauser & Yeung, 2010) and perceptual decision-making (O'Connell, Dockree, & Kelly, 2012) support the idea that these centroparietal signals, the late Pe and the P3b, reflect common neural computations related with the accumulation of internal evidences leading to the awareness of actions and events motivationally relevant for on-going behavior, such as performance errors. Therefore the frontal ERN and later centroparietal positive components seem to index different aspects of error processing: error-monitoring and error-awareness.

Recent ERP studies in trained musicians performing well-learned pieces or scales (Maidhof et al, 2010; Ruiz et al., 2009) revealed a frontal negative component similar to the ERN, occurring right before the onset of erroneous responses (occurring approximately 30-100 ms before the error). This “pre-ERN” component was interpreted as a prediction error signal at the level of motor preparation. This pre-

ERN signal is clearly distinguishable from other ERP responses related to movement preparation, such as the Readiness Potential (RP) (Shibasaki & Hallet, 2006) or the Lateralized RP (Gratton, Coles, Sirevaag, Eriksen, & Donchin, 1988) in terms of its spatial distribution and latency. The observation of error-related signals in the MFC even before an error was committed suggests that error-monitoring during overlearned performance is implemented through neural feed-forward computations. Still we do not have clear evidences how error-monitoring processes are engaged during the acquisition of musical skills.

Nevertheless, although the error-monitoring system has not been formally studied during the acquisition of rhythm skills, previous neuroimaging studies have shown that medial-frontal structures of the brain associated with action monitoring, such as the ACC and pre-SMA, are particularly active during early stages of motor skill acquisition (Floyer-Lea & Matthews, 2005; Jenkins et al., 1994; Jueptner et al., 1997; Petersen et al., 1998; Ramnani & Passingham, 2001; Toni et al., 1998). The greater activity of these regions early in learning, when the task is more demanding and errors are more likely, probably reflects an increased engagement of the error-monitoring system tracking inappropriate response tendencies, monitoring competition between multiple conflicting motor plans and, consequently, signaling the need for increased attentional control (Botvinick et al., 2001; Jueptner et al., 1997; Paus et al., 1998).

Drawing on this background, we designed a novel experiment in which musically naïve individuals learned to reproduce a series of rhythmic patterns across 12 trials of learning, by synchronizing tapping movements with auditory events with different time (rhythmic) intervals. We hypothesized that during early learning, when internal representations of the rhythms were weak there would be greater response conflict and thus greater demand for cognitive control leading to an increase of the amplitude in the ERN component (Botvinick et al., 2001; Gehring & Fencisik, 2001; Yeung, Botvinick, & Cohen, 2004). In contrast, we predicted that later in learning there would be a decrease of the ERN reflecting less engagement of the error-monitoring system tracking performance conflicts and recruiting control. Furthermore, we predicted that as learning progresses and participants develop stronger internal representations of the different templates associated to the learned rhythms, they will be more certain about their performance and errors would become more salient. We expected that during rhythm learning the accumulation of internal evidences leading to error-awareness should be accompanied by an increase of later centroparietal positive components.

## METHODS

### Participants

Eighteen students from the University of Barcelona (8 male, 10 female) between the ages of 19 and 31 (mean + SD = 21 ± 3 years old) participated in the study after giving informed written consent. All participants were right-handed, neurologically healthy and had normal hearing. Participants had no formal musical training besides the standard music classes at elementary school. Data from 3 participants were removed due to an excessive rejection rate of the EEG data (higher than 25%, see the methods below). Thus, 15 participants were included in the final behavioral and ERP analyses.

### Rhythm synchronization learning task

In this task participants first listened to and then tapped in synchrony with 8 different auditory rhythmic sequences. Each rhythm was presented 13 times in a single block so that learning could be assessed. Rhythms were delivered through stereo headphones (Creative HQ-1300) at a comfortable intensity level. Participants tapped in synchrony with the index finger of their right hand on a computer mouse. Stimuli were presented and responses recorded using Presentation software (Neurobehavioral Systems) on a PC computer. Response triggers were sent on-line to the PC recording the EEG.

The rhythms used in this experiment were based on those used in several previous studies (Chen, Penhune, & Zatorre, 2008b; Chen, Penhune, & Zatorre, 2008a). Each rhythm consisted of 11 woodblock notes, each 200 ms in duration, with total duration of 6 seconds. The 11 notes composing the rhythms had the same pitch, though the temporal organization of the notes differed such that different rhythm patterns were created. The interval following each sound (note) varied such that five different musical durations (onset-to-onset) were created. Each rhythm contained: five eight notes (each 250 ms), three quarter notes (each 500 ms), one dotted quarter note (750 ms), one half note (1000 ms) and one dotted half note (1500 ms).

The experiment was divided into 8 Blocks. In each block participants were trained on one of the 8 rhythms for 13 trials. Each learning trial included two conditions that always followed the same order: (1) *Listen* - participants were instructed to listen carefully to the rhythm without moving; and (2) *Synchronize* – participants were instructed to tap in synchrony with each sound in the rhythm and to avoid correcting their errors (Fig. 1). The beginning of each condition within each block was signaled by a warning tone 500 ms in advance. Between conditions there was 3.5 second pause for eye-blinking. Each block had duration of 9 minutes and on average the EEG session lasted 1.5 hours. Before the EEG session, participants were trained with 2 very basic rhythms to familiarize them with the task.

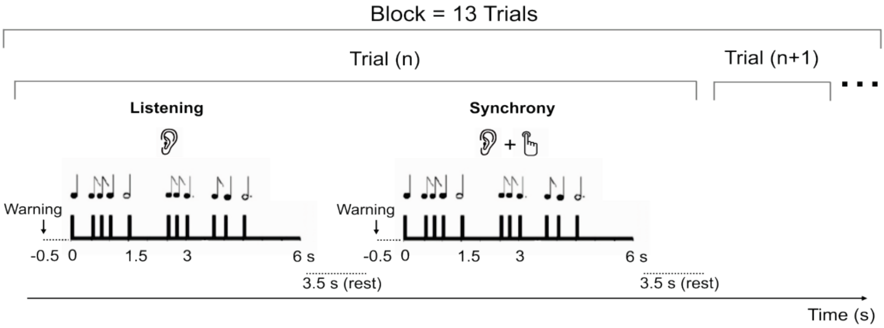


**Figure 1.** Schematic Diagram of the Task. Each trial comprises 2 different conditions that always follow the same order (Listening, Synchrony). After performing the 2 conditions participants move to the next trial. Each block includes 13 trials (the very first counts as a warm up trial). During one single block participants perform only one rhythm and the whole experiment involves 8 different rhythm sequences.

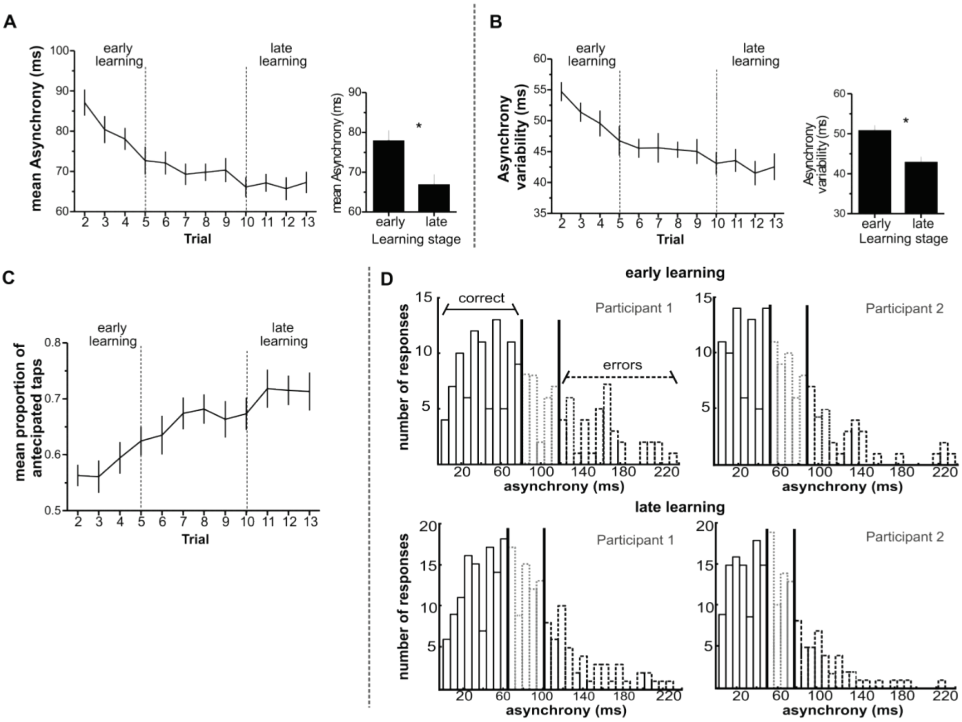


**Figure 2. (A)** Average of the mean asynchrony (absolute value, ms) pooled across the 8 rhythms during all trials (trials 2-13) and t-test comparisons of the mean asynchrony between *early Learning* (2-5 trials) and *late Learning* (10-13 trials) stages. **(B)** Same as **A** regarding the asynchrony variability. Error bars represent SEM. * p-value < 0.0001. **(C)** Average of the mean percentage of anticipated responses during the task (trials 2-13). **(D)** Graphic Representation of the method used to categorize error and correct taps in two random participants. For each individual, 50% of the total number of anticipated responses with small values of asynchrony were grouped as correct taps (solid black bars), and the 25% of the total number of responses with larger values of asynchrony was classified as error taps (dashed black bars).

### Behavioral analysis of rhythm synchronization performance

The first trial of each block was considered as a warm-up trial and excluded from the statistical analyses of both behavioral and ERP data. Rhythm production learning was assessed using three variables: (1) mean value of asynchrony (absolute value of the difference between the onset of each tap and the associated auditory tone in *ms*); (2) asynchrony variability (calculated as the SD of the asynchrony values), a measure related to performance consistency across learning  (Wolpert et al., 2011); and (3) mean proportion of anticipated responses (taps that precede the onset of the auditory tones).

For the calculation of the mean asynchrony and asynchrony variability we compared the onset of participant's taps to the onset of each tone of the rhythm. We examined both anticipated and delayed (taps following the auditory tones) taps. If more than one tap fell within the same time interval, the first was taken and the second was excluded (Chen et al., 2008a). Furthermore, if one tap appeared right before and another right after the stimulus onset we considered the one with smaller asynchrony. Based on the visualization of the distribution of the responses in all individuals, we decided to exclude from the behavioral and ERP analysis all taps with an asynchrony greater than 260 ms (absolute value) because they were very infrequent, less than 3% of the total number of responses. Asynchrony values were averaged for each trial of learning and pooled across rhythms to obtain an average across trials 2-13 of learning (12 learning trials in total). The same procedure was applied for the mean proportion of anticipated responses.

The evolution of participants' performance throughout learning was tested using one-way ANOVAS comparing each of the dependent measures across the 12 learning trials [mean asynchrony (Fig. 2A), asynchrony variability (Fig. 2B) and the proportion of anticipated responses (Fig. 2C)]. For the purposes of linking behavioral performance with changes in ERP measures, we divided the learning trials into two periods: *early learning* (trials 2-5) and *late learning* (trials 10-13). Differences in performance between the two learning periods were assessed with paired *t*-tests for each variable.

### Categorization of error and correct taps for the ERP analysis

During rhythm synchronization learning errors are relative, i.e. tap responses are more or less synchronous with the auditory stimuli. In this sense the perception of error commission relies on internal evaluations of the performer. As these evaluations may change over time with practice (while the performer consolidates his/her internal representations of the rhythm structure), the perception of error commission may follows the status of learning of the performer. For this reason, here we used a novel approach to categorize error responses in order to examine changes in EEG activity related to error-monitoring and error-awareness. Error and correct responses were categorized based on the distribution of the values of asynchrony of each participant's responses (anticipatory responses only, see the explanation below) during early (trials 2-5) and late (trials 10-13) learning stages separately. For each individual we took all his/her response asynchrony values and responses below the 50th percentile were categorized as correct (50% of responses with smaller values of asynchrony) and responses greater than the 75th percentile (the 25% of the responses with greater asynchrony) were categorized as errors (see examples from two representative participants at Fig. 2D).

Further, the number of anticipated error responses were matched across early and late learning conditions, such that there weren't differences between the number of errors in early learning (mean + SD = 45 + 5.6; minimum: 37; maximum: 55) and late learning stages (mean + SD= 46 + 6; minimum: 38; maximum: 58) ($t(14)=1.4$, $p >0.05$) that could explain potential differences in the ERP results.

We chose to analyze only anticipatory responses because they were more common (65%) than the delayed responses. Besides, it was observed that after delayed responses the onset of ERP components associated with fast error processing mechanisms (e.g. ERN/Pe components) overlapped in time with ERP signals related with the processing of the auditory tones (e.g. the auditory N1 component). Thus, using only anticipated responses avoided confounds in the interpretation of the ERP data.

### EEG recording and data analysis

EEG was recorded from tin electrodes mounted in an elastic cap located at 29 standard positions (Fp1/2, Fz, F7/8, F3/4, Fc1/2 Fc5/6, Fcz, Cz, C3/4, T3/4, Cp1/2, Cp5/6, Pz, P3/4, T5/6, PO1/2, Oz). Vertical eye movements were monitored with an electrode at the infraorbital ridge of the right eye. While performing the rhythms participants were instructed to fixate their eyes on a black square in the middle of the screen (grey background) in order to avoid eye movements. Between the listen and synchronize conditions there was 3.5 second pause for eye-blinking. Electrode impedances were kept below 5 kΩ. The electrophysiological signals were digitized at a rate of 250 Hz and filtered with online band-pass of 0.01–70 Hz (half-amplitude cutoffs). All scalp electrodes were referenced offline to the mean activity of the left and right mastoids.

Epochs of 1000 ms before and 1000 ms after the motor response (tap) were extracted from the EEG and baseline was corrected from -200 to -50 *ms* prior the response onset (response-locked ERP analysis). ERPs associated with the processing of the auditory tones composing the rhythms were also computed on epochs of 700 ms starting 100 ms before the onset of auditory tones; baseline was calculated from -50 to 50 ms to minimize misalignments of the waveforms based on anticipatory neural activity (Lange, 2011) and to overcome problems in baseline shifts due to the

ERP deflections following the motor response. Trials exceeding ± 80 μV in both EEG and EOG during the epoch window were rejected offline. Only epochs that were preceded by at least 800 *ms* of error-free responses were entered in the analysis (Ruiz et al., 2009). EEG was offline low-pass filtered at <14 Hz for both response-locked and stimulus-locked ERP analysis. The data was further filtered with a band pass filter [3-9 Hz] in order to isolate theta-band specific ERPs associated with error and conflict monitoring (Cavanagh et al., 2012; Luu & Tucker, 2001). This method is also valuable to remove positive slow wave potentials in which error-related negative ERPs developed and stabilize possible drifts.

### *Analysis of ERP components locked to the motor response*

The analysis of the grand-average ERPs following the motor response showed that error taps were followed by a cascade of different ERP components that developed during four distinct time windows (Fig. 3). Right after the response onset an ERN was observed in error taps. The ERN was followed by the early Pe component (Arbel & Donchin, 2009; O'Connell et al., 2007). In turn, the ERN and the early Pe components gave rise to another two ERP deflections, respectively labeled as the N1/Auditory-feedback Negativity (AFN) component and the centroparietal P3 component. As the N1/AFN developed around the onset of the auditory tones composing the rhythm sequences, we questioned whether this component would be linked to auditory evoked activity, actually representing the standard auditory N1 component (Hillyard, Hink, Schwent, & Picton, 1973; Woldorff et al., 1993) and, therefore, dissociated from the motor output process (i.e. ERN/Pe). To answer to this question single-trial ERPs analysis was additionally conducted (see for a similar approach Burle et al., 2008; Delorme & Makeig, 2004; Jung et al., 2001) as well as stimulus-locked ERP analysis (see below).

The inspection of single-trial ERPs was also critical to validate our approach to categorize correct and error taps. This analysis allowed us to inspect the EEG signal without averaging and, therefore, examine transient changes of the signal as function of the asynchrony variability (our measure to discriminate error from correct taps in all participants) (Burle et al., 2008; Jung et al., 2001).

Single-trial ERP epochs were analyzed from the EEG filtered data (low-pass <14 Hz) in the electrode Fcz (fronto-central scalp location), which is commonly used in ERP analysis of error processing. In Figure 4 we depict single-trial ERPs as a color-coded map in which color-changing values reflect the intensity of the EEG signal recorded at a certain electrode location (Fcz), for a given trial and time point. Single-trials were sorted by decreasing values of asynchrony (*y*-axis), in order to explore ERP differences between small and larger values of asynchrony and displayed after smoothing with a narrow of 200 points moving window to increase the salience of stimulus and response locked features (Jung et al., 2001). Differences in the intensity of the signal are represented as dark blue (negative voltage polarity) and yellow-orange (positive voltage polarity) color activations. In all plots the response onset time (*grey vertical line)* and the absolute value of asynchrony at all single trials, that is, the moment of the auditory feedback presentation (*oblique black line*) are represented.

Time-windows for the statistical analysis of response-locked ERPs were chosen by the visual inspection of the grand-average waveforms (Fig. 3 and Fig. 5). Four time windows were selected: (0-50 ms, ERN); (90-150 ms, Pe); (180-250 ms, N1/AFN); and (300-450 ms, P3). The mean voltage of these four ERP components was subjected to ANOVAs with Type of Response (error, correct), Learning Stage (early learning, late learning) and Electrode Location (frontral, Fz; fronto-central, Fcz; and central, Cz) as within-subject factors. The Greenhouse-Geisser epsilon was used to correct possible violations of the sphericity assumption. P-value after the correction is reported.

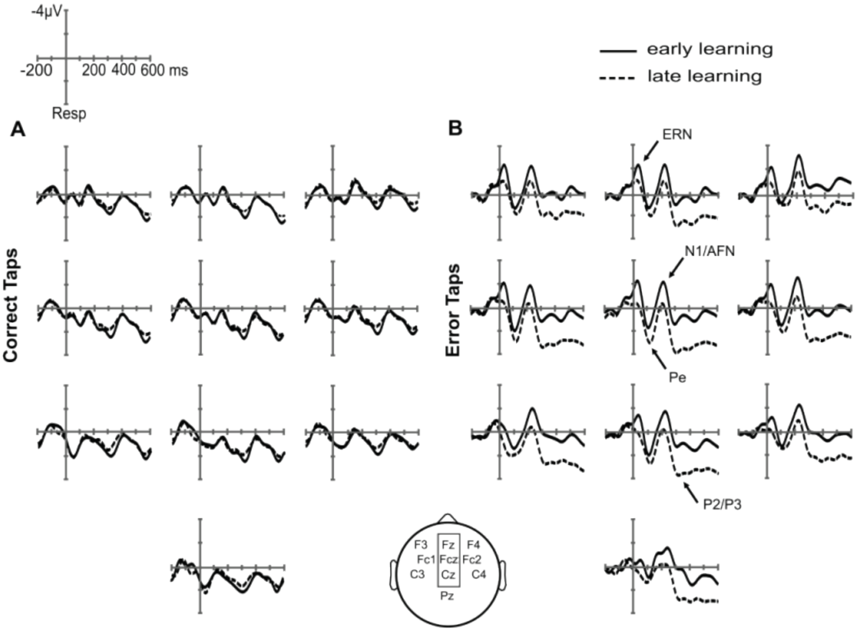


**Figure 3.** ERPs Response-locked to the motor responses made in anticipation to the auditory tones during *early learning* (solid line) and *late learning* (dashed line) stages for correct (left panel, **A**) and error (right panel, **B**) taps.

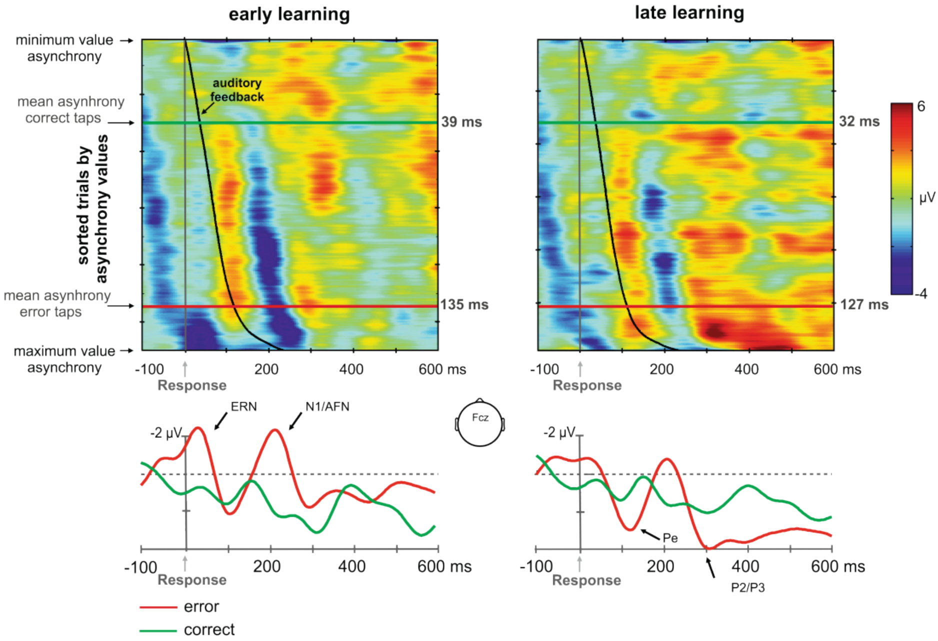


**Figure 4.** Representation of single-trial ERPs containing all trials (all participants) at Fcz electrode for the *early* and *late learning* blocks. In the y-axis are represented single trials sorted by the absolute value of asynchrony, and in the x-axis is represented the time respect to the response onset. The color represents the intensity of the signal for each time point and trial. The response onset time is indicated by the *grey vertical line* and the absolute value of asynchrony for a single trial (auditory feedback onset) is represented by the *oblique black line*. The mean value of asynchrony for error (*red line*) and correct (*green line*) taps in both learning stages is also indicated. In the bottom of the figure are plotted the grand average response-locked ERPs of all error and correct taps.

### *Analysis of ERP components locked to the auditory feedback*

To tease apart the possibility that the error-related ERP effects observed for the N1/AFN and P3 components during the response-locked analysis were due to baseline problems resulting from the ERN/Pe deflection, stimulus-locked analysis were also conducted on error and correct taps for both learning periods (note that taps always preceded the auditory tones). Moreover this analysis allowed us to confirm that the onset of these two ERPs always followed the auditory tones composing the rhythms (i.e. auditory feedback). ANOVAs with variables Type of Response (error, correct), Learning Stage (early learning, late learning) and Electrode Location (frontral, Fz; frontal-central, Fcz; and central, Cz) were computed on the mean voltage of these two components that developed following the onset of the auditory tones: the N1/AFN (80-120 ms) and the centroparietal P2/P3 (160-300 ms). Note that both the N1 and the P2/P3 signals from the stimulus-locked analysis reflect, respectively, the response-locked N1/AFN and P3 components. For heuristic proposes we labeled the ERPs from the stimulus-locked analysis as N1/AFN and centroparietal P2/P3 components. The Greenhouse-Geisser epsilon was used to correct possible violations of the sphericity assumption and P-value after the correction is reported.

### Correlation analysis between behavior and ERP data

Furthermore we examined whether performance consistency, which was measured by the mean asynchrony and asynchrony variability throughout all learning trials, was correlated with neurophysiological responses that have been previously associated to error awareness, i.e. enhancement of centroparietal positive P3-like components (Murphy et al., 2012; Overbeek et al., 2005; Steinhauser & Yeung, 2010). We reasoned that more consistent or accurate performance might indirectly asses the accuracy of the internal representations built in relation to the target rhythm and that an increase in consistency might be related to larger P3 amplitude, the component associated to error awareness. To test this hypothesis, Pearson correlations between the mean value of asynchrony and asynchrony variability across all trials and the amplitude of the centroparietal P3 response (calculated by the difference error - correct taps across early and late learning stage) were computed.

## RESULTS

### Behavioral performance during rhythm learning

Behavioral measures of performance showed significant improvements across the 12 trials of learning (Fig. 2) with decreases in mean asynchrony ($F(1,14) = 22.1$, $p < 0.001$); asynchrony variability ($F(1,14) = 55.2$, $p < 0.001$) and mean proportion anticipated responses ($F(1,14) = 55.9$, $p < 0.001$). For the purposes of comparison with the EEG data, learning was divided into early (trials 2-5) and late (trials 10-13) phases. Consistent with the results across all trials, comparison of performance across the two phases showed significant improvements for all three measures [mean asynchrony: early learning (mean ± SD), 77.8 ± 11 ms; late learning, 66.8 ± 10 ms: $t(14) = 4.8$, $p < 0.001$; asynchrony variability: early learning (mean ± SD), 50.6 ± 6 ms; late learning, 42.7 ± 6 ms: $t(14) = 6.4$, $p < 0.001$; mean proportion of anticipated responses: early learning (mean ± SD), 0.6 ± 0.1; late learning, 0.72 ± 0.1: $t(14) = 7.4$, $p < 0.001$].

The behavioral results showed that throughout the learning trials participants benefit from practice and improved their skill to tap more synchronized with the auditory tones composing the musical rhythms and with fewer fluctuations.

### EEG signals of error-monitoring during rhythm learning

#### *Error-monitoring based on the motor response (response-locked ERPs)*

The inspection of the grand-average ERPs following the motor response revealed that error taps led to a cascade of four distinct ERP components (ERN, Pe, N1/AFN and P3) which were not observed in correct taps (Fig. 3).

The ERN was visible immediately after the production of an error (Fig. 3B and 5), peaking at about 0-50 ms. The ERN was also seen on the single-trial ERP analysis (Fig. 4) as a negative EEG signal starting right before the onset of the motor response (vertical grey line) and reaching maximal activity right after the response onset (notice the ERN peak in trials with larger values of asynchrony, as signaled by the red line representing the mean value of asynchrony in error taps). Typically, for the comparison between error and correct taps during early and late learning at the three midline recording sites (Fz, Fcz and Cz), the ERN was increased for error compared to correct taps (main effect of Response Type ($F(1,14) = 60$, $p < 0.001$) (Fig. 5A). The ERN showed a frontocentral topographical distribution with a maximum at the Fz site (Fig. 5C). An interaction Response Type x Learning ($F(1,14) = 7.8$, $p < 0.05$) revealed that the ERN following errors was larger during early stages of rhythm learning when compared to the later learning period across all three recording sites (Fig. 3B and 5A; see also the single-trial analysis, Fig. 4). Next, the data was reanalysed with a band pass filter [3-9 Hz] in order to isolate error-related theta-band specific ERPs (Cavanagh et al., 2012; Luu & Tucker, 2001) (Fig. 5B). Again it was observed a greater ERN after error taps (main effect of Response Type: $F(1,14) = 29.2$, $p < 0.0001$) and especially during early stages of rhythm learning (Response Type x Learning interaction: ($F(1,14) = 9$, $p < 0.01$).

Following the ERN, around 90-150 ms, the Pe component was identified as an early positive deflection that was enhanced after error taps (main effect of Response Type: $F(1,14) = 6.2$, $p < 0.05$) (Fig. 3B and 5A). The early onset of this Pe and its frontocentral distribution (see on Fig. 5C the spatial map of this ERP) differentiates this Pe from the later centroparietal Pe which has been recently associated with error salience and consequent error-awareness (Arbel & Donchin, 2009; Murphy et al., 2012; O'Connell et al., 2007; Steinhauser & Yeung, 2010). Regarding the early Pe enhancement after error taps, no main effect of Learning ($F(1,14) = 1.6$, $p > 0.05$) or its interaction with Type of Response ($F(1,14) = 3.6$, $p = 0.078$) was observed over the three recording sites.

Following the ERN/Pe components a second negative component was shown. As it can be seen in the Figure 4 (single-trial ERP analysis) this negativity, peaking at about 180-250 ms after response onset, developed along with the presentation of the auditory tones (highlighted by a dark oblique line). The computation of single-trial ERPs was critical to differentiate this ERP from the ERPs related with the motor response (ERN/Pe deflections). Note that this negativity appears ~ 100 ms after the auditory tones onset, which in terms of latency is very similar to the auditory N1 component (Hillyard et al., 1973; Woldorff et al., 1993). This negative component, termed N1/AFN, was increased in error taps compared with correct taps (main effect of Response Type: $F(1,14) = 77.0$, $p < 0.0001$) (Fig. 5a) with a maximum at Fcz electrode (Response Type x Electrode interaction: $F(2,28) = 4.0$, $p < 0.05$) (see the fronto-central topographical distribution of the AFN on Fig. 5C). No Learning effect was observed regarding this negative component ($F(1,14) < 1$) or the interaction Response Type x Learning ($F < 1$). After applying a band pass filter in the theta band, which was crucial to remove the positive wave which this component developed (the early Pe), we confirmed the anterior findings by showing that the enhancement of the N1/AFN response after errors was not modulated by learning (both main effect of Learning and the interaction Type of Response x Learning were not significant ($F < 1$) (Fig. 5B).

Finally, both the grand-average waveforms and single-trial ERP analysis revealed that error taps during the late learning stage were associated with the increase of a later positive ERP occurring 300-500 ms after the error onset (the centroparietal P3). By looking carefully at Figure 4 it can be seen that this positive ERP was greater in responses with larger values of asynchrony. This later positive response, representing the P3 component, showed a centroparietal distribution with a maximum at Cz electrode (Fig. 5C) and was marginally modulated by learning (main effect of Learning: $F(1,14) = 4.1$, $p = 0.06$). An interaction Response Type x Learning ($F(1,14) = 19.5$, $p < 0.001$) showed that this enhanced later centroparietal P3 for error taps was specific for later stages of learning (Fig. 3A and 5A).

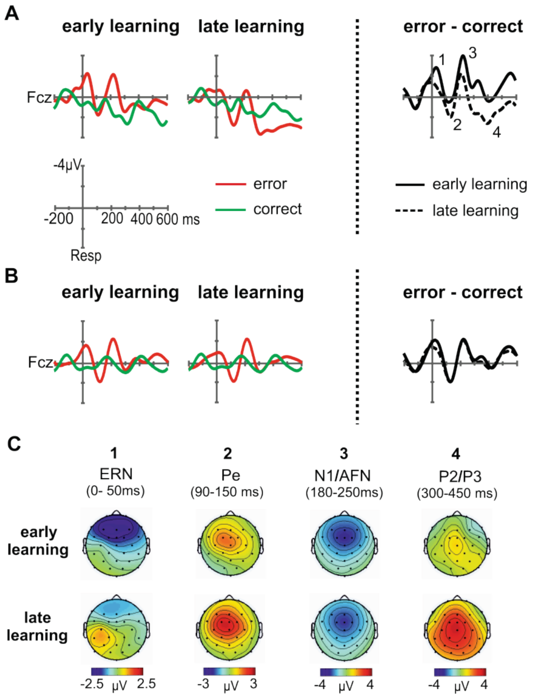


**Figure 5.** ERPs Response-locked to error (red) and correct (green) taps and for the difference waveform (error-correct) during *early learning* (solid line) and *late learning* (dashed line). **(A)** Grand averages low-pass filtered < 14 Hz. **(B)** Grand averages band pass filtered (Theta range: 3-9 Hz). **(C)** Topographical maps of the frontal ERN (1), the frontocentral Pe (2); the N1/AFN (3) and the centroparietal P2/P3 (4).

### *Error-monitoring based on auditory feedback processing (stimulus-locked ERPs)*

The results from the stimulus-locked analysis (Fig. 6) replicated the findings regarding the N1/AFN and P3 modulations from the ERP analysis timed to the motor response.

Error taps compared with correct taps elicited an increased N1/AFN with maximal activation at Fcz location [main effect of Response Type: $F(1,14) = 14.2$, $p < 0.005$; and Response Type x Electrode interaction: $F(1,14) = 16.8$, $p < 0.001$]. The N1/AFN peaked at about 100 ms after auditory stimulus onset as the classical N1 component. Note that the topographical distribution of the N1/AFN from the stimulus-locked analysis and the N1/AFN (response-locked analysis) after error taps is very similar; therefore the two ERPs may reflect the same neural mechanism. Alike the N1/AFN results from the response-locked analysis, no Learning effect ($F(1,14) = 2.9$, $p > 0.05$) or the interaction between Response Type x Learning ($F(1,14) = 2.7$, $p > 0.05$) was shown for the stimulus-locked N1/AFN component.

Moreover, the inspection of the ERPs associated with auditory feedback processing reconfirmed our previous result stating the specific increase of the centroparietal P3 component after error taps during the late learning stage. In error taps, after the N1/AFN response we observed a later centroparietal positive component which was modulated by Leaning ($F(1,14) = 6.3$, $p < 0.05$) and was enhanced during the late learning stage in all recording sites (Response Type x Learning ($F(1,14) = 17.1$, $p < 0.005$). Notice that this centroparietal positivity (labelled as P2/P3 response) reflects the same component in both the response-locked and stimulus-locked cases as can be clearly observed when comparing their scalp distributions (see Fig. 5C and Fig. 6B).

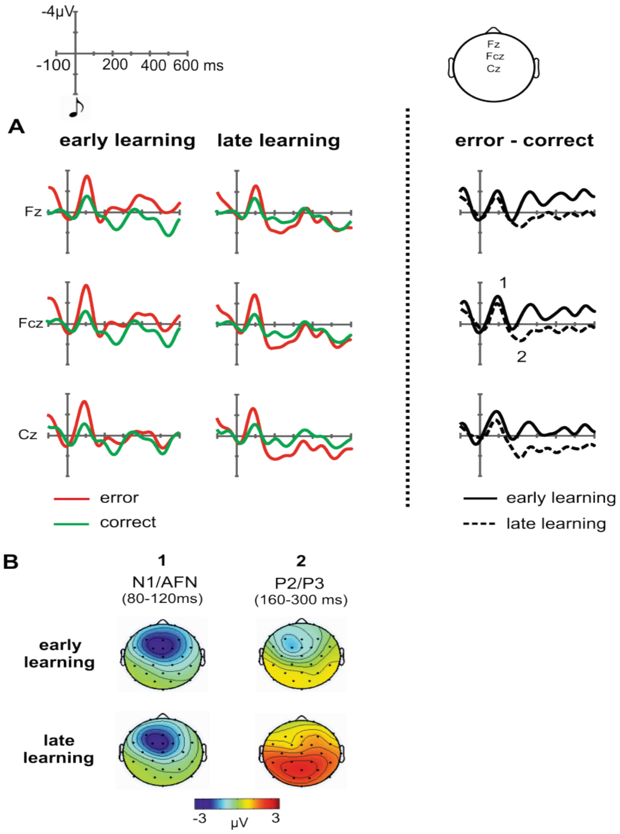


**Figure 6. (A)** ERPs Stimulus-locked to the auditory feedback (i.e. tones composing the rhythms) for error (red) and correct (green) taps and the difference waveform (error-correct) during *early learning* (solid line) and *late learning* (dashed line). **(B)** Topographical maps of error-related activity of the N1/AFN and centroparietal P2/P3components.

### Correlation analysis between performance accuracy and the late centroparietal positive component

We next examined whether whether individual performance consistency, which was measured by the mean asynchrony and asynchrony variability throughout all learning trials was associated with participants degree of error awareness, measured by enhanced activity in later centroparietal positive components. Participants with larger centroparietal P3 amplitude after erroneous taps reproduced the rhythmic sequences more synchronized with the tones (smaller values of asynchrony; $r = -0.66$, $p = 0.007$) and with less fluctuations (smaller values of asynchrony variability: $r = -0.58$, $p = 0.022$) during the whole task.

## DISCUSSION

In the present study, we tracked ERP changes associated with error-monitoring and error-awareness as musical untrained individuals learned to reproduce series of rhythmic patterns. Behavioural findings showed that across the learning trials participants improved in their ability to reproduce the rhythms and to synchronize their motor responses with the auditory stimuli. Changes in performance were paralleled by changes in cortical brain activity related to error-monitoring and error-awareness. The ERN amplitude was greater in early learning when compared to late learning, which indicates a greater involvement of the error-monitoring system when performance was more demanding and more cognitive control was required. It was also observed that in both stages of learning, erroneous compared to correct taps led to larger auditory evoked responses, indexed by the amplitude of the N1/AFN component, suggesting automatic attention orientation to the auditory tones after errors in performance. In agreement with our predictions, later in learning, errors were followed by a centroparietal P3-like component (the P2/P3) that was not present in the initial stages of learning. Importantly, we showed that the magnitude of this late centroparietal positive ERP was associated with individual differences in tap synchrony and asynchrony variability, suggesting a relationship between error-awareness and more accurate performance.

### Error-monitoring and control mechanisms during early stages of rhythm learning

A key finding of our study was that the ERN was modulated by learning, showing greater amplitude in early compared to late learning. The ERN is hypothesized to represent a neural marker of error and conflict monitoring (Gehring et al., 1993; Holroyd & Coles, 2002; Yeung et al., 2004). Many studies have pointed the MFC, specifically the ACC and pre-SMA, as the source of the ERN (Debener et al., 2005; Dehaene et al., 1994; Yeung et al., 2004). The MFC is a key region for action monitoring and regulative aspects of motor and cognitive control, being critically involved in the flexible optimization of behavior and learning (Botvinick et al., 2004; Holroyd & Coles, 2002; Ridderinkhof et al., 2004; Ullsperger, Danielmeier, & Jocham, 2014).

The larger amplitude of the ERN during early stages of rhythm learning might reflect greater error-monitoring and response conflict experienced by the participants (Carter et al., 1998; Gehring & Fencisik, 2001; Yeung et al., 2004) as they tried to accurately reproduce the rhythm patterns when the representation of the appropriate response was as yet underdetermined (Botvinick et al., 2001). This finding points to a greater engagement of the MFC monitoring and conflict detection system (Botvinick et al., 2004; Nachev, Wydell, O'Neill, Hussain, & Kennard, 2007; Taylor, Nobre, & Rushworth, 2007) when performance was more challenging and more cognitive control and attention to the task was needed (Botvinick et al., 2001; Bush et al., 1998; Paus et al., 1998; Ramnani & Passingham, 2001). Early in learning the ERN may function as a signal to increase higher-level cognitive control to redirect attention to auditory feedback in order to overcome conflict among multiple possible motor plans (Botvinick et al., 2001; Kerns et al., 2004; Ridderinkhof et al., 2004). Conversely, as learning progresses and performance becomes less effortful and more automatized, the error-monitoring signal decreases. This decrease in error-monitoring might as well be associated to a diminished need to implement fast and very subtle error-correction processes, as less partial or full erroneous responses might exist. Thus, this decrease in error-monitoring activity might be associated to a reduced reliance on top-down control systems once performance becomes more fluent.

Considering that the ERN component has been associated with conflict and error processing in the MFC (Debener et al., 2005; Ridderinckoff et al., 2004; Holroyd & Coles, 2002; Botvinick et al., 2001) our results are in agreement with previous neuroimaging findings showing greater activation in medial prefrontal regions (specifically the ACC and pre-SMA) during the early stages of motor skill acquisition (Floyer-Lea & Matthews, 2005; Jenkins et al., 1994; Jueptner et al., 1997; Petersen et al., 1998; Ramnani & Passingham, 2001; Toni et al., 1998). Although error and conflict monitoring was not formally examined in those studies, it is very likely that the observed activations in the MFC were related to neural computations involved in the monitoring and compensation of a larger number of conflictive and error responses during the first stages of learning. Similarly, Brown and colleagues (Brown et al., 2013) have recently evaluated pianists learning novel melodies. Improvements in performance with repetition were accompanied in this study by linear BOLD decreases in both the ACC and pre-SMA, which may indicate conflict resolution between multiple motor plans (*see* Nachev et al., 2007). In another study and using a different behavioral task (the Stroop task), it was observed that ACC activity related with response conflict diminished with practice (Bush et al., 1998). Importantly also, from animal studies, it has been observed that the ACC is a crucial region in rapid associative learning (Gabriel, 2002) and sequencing learning (Nakamura, Sakai, & Hikosaka, 1998; Procyk, Tanaka, & Joseph, 2000).

Moreover, across both stages of learning, error taps were followed by larger auditory evoked responses (indexed by the amplitude of the the N1/AFN component) compared to correct taps. Changes in the amplitude of the N1 could be associated with an attention orientation mechanism by which sensory processing in the auditory cortex is modulated (Hillyard et al., 1973; Woldorff et al., 1993). This finding is consistent with a previous study of audiomotor synchronization that also showed the modulation of the auditory N1 component in response to perturbations in the temporal organization of the auditory feedback and the resulting motor corrections (Praamstra, Turgeon, Hesse, Wing, & Perryer, 2003).

Although speculative, it is probable that following errors the error-monitoring system may trigger changes in top-down attentional and motor control mechanisms that in turn modulate task-relevant sensory areas to promote post-error adaptations. The enhancement of auditory processing after errors is in line with recent fMRI evidences showing the modulation of task-relevant perceptual areas after errors via error-related medial frontal activity (Danielmeier et al., 2011; King et al., 2010). Importantly, however, our data indicate that attention orientation after errors is not modulated by learning. Thus for novice performers external feedback may be useful for post-error adjustments during all stages of skill acquisition. Although it has been proposed that auditory feedback is not crucial for error-monitoring in highly skilled performers (Maidhof et al., 2010; Ruiz et al., 2009), other evidence indicates that when auditory feedback is present performers still rely on the comparison between the auditory input and the motor output (Finney & Palmer, 2003; Pfordresher & Palmer, 2006; Pfordresher, 2006).

### Error-awareness depends on the formation of internal templates of the target rhythm

The present study also revealed that during later learning errors were followed by a large centroparietal positive component (which was labeled as the P2/P3 component) appearing after the auditory feedback and which was not present early in learning. The latency and the topographical distribution of this centroparietal positive ERP resembles the late Pe and the P3b component (Arbel & Donchin, 2009; Ridderinkhoff et al., 2009), which have been recently related to brain mechanisms associated to the conscious perception of errors or salient events in performance (Murphy et al., 2012; O'Connell et al., 2007; Ridderinkhof et al., 2009; Steinhauser & Yeung, 2010; Ullsperger, Fischer, Nigbur, & Endrass, 2014). Importantly, we did not observe any modulations of the early frontocentral Pe with learning that could suggest an association of this component with the subjective experience error awareness.

Although errors are very salient events, during early stages of learning error detection is accompanied by some degree of uncertainty. The increase of the centroparietal P3 late in learning may be evidence that an internal template of the rhythm structure has been established in working memory (Donchin & Coles, 1988) increasing participants' capacity to evaluate their response accuracy. In this sense, our findings are consistent with the idea that centroparietal P3-like signals reflect neural computations underlying the accumulation of perceptual evidences leading to the detection of salient events to performance (Hillyard et al., 1971; O'Connell et al., 2012) and, therefore, may be associated to the strengthening of internal evidence leading to error awareness (Steinhauser& Yeung, 2010).

Previous EEG studies on sequence learning have also described an association between centroparietal P3 signals and explicit knowledge about task rules (Baldwin & Kutas, 1997; Ferdinand, Mecklinger, & Kray, 2008). Further, professional musicians and conductors, who likely have strong internal models of expected sounds also exhibit larger amplitude of the P3 in response to deviant auditory stimuli when compared to music naïve subjects (Nager, Kholmetz, Altenmuller, Rodriguez-Fornells, & Munte, 2003). The present results are also in agreement with previous studies showing the relevance of the P3 on the detection of violations in sequential patterns (Brochard et al., 2003; Ford & Hillyard, 1981, Jongsma et al., 2004). For example, Jongsma and colleagues (Jongsma, Meeuwissen, Vos, & Maes, 2007) showed that the amplitude of the P3 component was sensitive to the ability of their participants to detect small tempo changes. Importantly, we showed that the magnitude of the centroparietal P3 was associated with individual differences in tap synchrony and asynchrony variability, suggesting a relationship between improved performance and error-awareness. These findings give empirical support to the idea that later centroparietal positive ERPs are positively related with learning on the basis of trial-and-error (Overbeek et al., 2005). In contrast, no relationship was found between the accurate performance and the amplitude of the ERN and performance. This is consistent with the idea that the ERN and later centroparietal positive ERPs reflect different aspects of error processing (Overbeek et al., 2005; Steinhauser & Yeung, 2010) that are differently affected by learning. Therefore, while the ERN may index more general and automatic monitoring processes (Nieuhuinhuis et al., 2001) such as computing the degree of response conflict (Yeung et al., 2004), later centroparietal P3 signals may tap the subjective experience of error awareness.

## CONCLUSION

The present results shed new light on how electrophysiological responses associated with error-monitoring and error-awareness are modified during the acquisition of new skills. Consistent with previous research we found that error-monitoring and error-awareness are partially dissociable processes in the human brain. During skill acquisition, error-monitoring processes indexed by the ERN were more engaged early in learning when the internal representation of the target response was underdetermined, conflict between competing motor responses was higher and greater recruitment of attentional control mechanisms was required. In contrast, error-awareness, identified by a late centroparietal P3 component, emerged only during later stages of learning, when the representation of the target rhythm was stronger and errors become more salient. Further, we extended previous findings by showing that during skill acquisition error-awareness is a byproduct of learning that likely relies on the consolidation of internal templates of the task and which is related to performance accuracy. Finally, the enhancement of the N1/AFN component after errors was consistent across both early and late learning, suggesting that in naïve participants errors trigger engagement of auditory attentional mechanisms that are automatic and independent of learning. We hypothesize that enhanced attention to auditory feedback following errors contributes to both error correction and the strengthening of the representation of the learned response. Overall, we think that this study is important in extending previous work on error processing mechanisms in simple reaction time tasks to more ecologically valid contexts involving learning of new skills.

# 5

# Study 3

## Violating action semantics: neural signatures of self-generated and alien-errors [*]

---

[*] This study corresponds to:
Padrão, G., Gonzalez-Franco, M., Sanchez-Vives, M.V., Slater, M., Rodríguez-Fornells, A. Violating action semantics: neural signatures of self-generated and alien-errors. (**submitted**)

## INTRODUCTION

In normal circumstances the actions of our own body are attributed to our agency, including the errors we make. This is the ordinary experience of ourselves, where we typically experience an implicit and diffuse sense of control over our actions (Gallagher, 2000). In the present study, though, we investigated whether one's feeling of agency would be affected when our surrogate body carries out actions that were not intended by us in order to disambiguate the neural processes underlying the monitoring of self-generated errors and errors that are caused without subjects' will and which may undermine one's natural sense of agency.

Previous EEG studies in humans have identified event-related brain potential (ERP) correlates underlying the monitoring of erroneous actions. The so-called Error-related negativity (ERN) is a frontocentral ERP originated in the medial-prefrontal cortex that is elicited by self-generated errors, negative outcomes and error observation (Gehring et al., 1993;Holroyd et al., 2005;van Schie et al., 2004). It is thought that the ERN provides a neural index of an *internal error-monitoring system (internal-EMS)* which evaluates the correctness of on-going actions and predicts fast motor corrections and compensatory behaviors when erroneous response tendencies are detected (Danielmeier et al., 2011;Marco-Pallares, 2008;Rodriguez-Fornells et al., 2002). However, the extent to which the neural processes underlying the ERN are also invoked by situations wherein subjects partially loss their sense of control has not been explored yet.

It has been suggested that our sense of agency relies on an retrospective error-monitoring process (referred to here as the *external error-monitoring system, external-EMS*) that determines whether the predicted consequences of an on-going action (based on internal "efference"copies of the motor command) are coherent with the sensory feedback information resulting from the actual performed action (Frith et al., 2000;Jeannerod, 2006). Neuroimaging studies suggest that manipulating the correspondence between an intended action and the actual consequences of a movement (ex. by changing the spatial or temporal contiguity of that movement) affects subject's sense of agency and leads to increased activity of the inferior parietal cortex (Farrer et al., 2003;Farrer et al., 2008b). However, no previous study has examined a putative neurophysiological response of this external-EMS as well as the time required by the human brain to compute agency violations during self-generated movements.

The objective of the present research has been to functionally dissociate the neurophysiological mechanisms underlying the internal and the external error-monitoring systems. To accomplish this goal we performed two ERP experiments. In the first experiment we set up a novel setting in which we recorded ERPs of healthy participants embodied into avatars while they carried out an error-prone reaction time task immersed in a virtual reality environment. Critically, in few occasions participants' correct responses were falsified by an "erroneous" movement of their embodied avatar (i.e. alien-errors) which perturbed their sense of agency. ERP signals related to self-generated errors and (false)-alien-errors were then compared . In a second ERP experiment we examined whether ERP signatures related to agency violations resulting from alien-errors could be attributed to processes associated with error observation.

## METHODS

### Participants

Eighteen neurologically healthy right-handed volunteers from the Faculty of Psychology at the University of Barcelona participated in the main ERP experiment (Exp. 1) (6 men; mean age, 26 ± 7 years). Two weeks later nine participants from Exp. 1 (3 men; mean age, 25 ± 8 years) returned to the lab to participate in a control ERP experiment (Exp. 2). All of them gave informed consent according to the declaration of Helsinki and were paid for their participation. The ethical committee from the University of Barcelona gave approval to the present project.

### Procedure and experimental task design

#### *Procedure*

Participants performed a standard error-prone reaction time task, the Eriksen Flanker Task, augmented by a whole body illusion in an Immersive Virtual Environment (Fig. 1, see also supplementary movie S1).

Each participant's own body was substituted by a gender matched avatar from a first person perspective (1PP) such that when participants looked down at themselves they saw an avatar body replacing their own. Through providing correct multisensory correlations and 1PP of a virtual body it is possible to induce whole body ownership illusions "It felt as if the virtual body and its movements were my body and my movements" (Lenggenhager et al., 2007;Petkova and Ehrsson, 2008;Slater et al., 2010). These experiences are similar to body illusions evoked by synchronous multisensory stimulation in which we can incorporate into our body schema the body parts of others or objects that are not part of our body (Botvinick and Cohen, 1998;Ehrsson et al., 2004;Sanchez-Vives et al., 2010)

Participants were fitted with a stereo NVIS nVisor SX111 head-mounted display (HMD). The HMD has dual SXGA displays with 76°H x 64°V degrees field of view (FOV) per eye, totalling a wide field-of-view 111° horizontal and 60° vertical, with a resolution of 1280 x 1024 per eye displayed at 60Hz. Head tracking was computed by a 6-DOF Intersense IS-900 device. Participants' movements were tracked using an optical infrared system (12 camera OptiTrack). Whole arm kinematics (hand, elbow and shoulder positions and rotations) were computed from the hand position using Inverse Kinematics. Our setup supported the real-time display of the avatar with 3 degrees of freedom (DOF) in the head and 4 DOF in the right arm giving the participant strong visual-motor coherence and an illusion of embodying the avatar. The virtual environment was programmed in the XVR system (Tecchia et al., 2010) and the virtual character displayed through the HALCA library (Gillies & Spanlang, 2010).

In the virtual version of the Eriksen flanker task stimuli consisted of three black arrows oriented horizontally, one central (target) and two flanker arrowheads above and below (Fig. 1b). The trials were categorized as compatible, when the central arrow (target) had the same direction of flankers, or incompatible, when target and flankers had opposite directions. The proportion of incompatible trials (60%) was greater than the proportion compatible trials in order to increase the number of self-generated errors. The duration of each stimuli presentation was 150 ms and the interval between two successive presentations (SOA) was 1150 ms; a fixation cross was presented during this interval, 1000 ms, after the disappearance of the stimulus array. Participants were instructed to respond both accurately and quickly to the direction of left or right-target pointing arrows by moving their right hand in the direction of two assigned buttons that were located on the right and left sides of the table (the buttons of response were located both on the displayed virtual table and on the real table, so the visual-tactile contingencies were coherent) (Fig. 1a and 1b). Participants were instructed to give only one response per trial and to avoid error-correcting movements. At each trial after the response was performed the hand should return to the initial position (0,0,0), equidistant to the two buttons. The recorded hand-tracked movements were analysed offline in order to compute the time of movement onset in each trial and, consequently, evaluate response accuracy and analyze response-locked ERPs. To detect the participants' responses we used the projected position of the hand over the x axis – as the Euclidean distance between the two response buttons. For each trial we distinguished among four different response types: correct, error, error-corrected and not responded. Once the response was classified we used the derivative of the position to assess the onset movement described as the moment just before the hand starts to move (see Fig. 2).

### *Experimental design*

The main ERP experiment (Exp. 1) was divided in two bloc conditions, *Congruent Movement (CM) condition* and the *InCongruent Movement (InCM) condition*. The experiment started with the CM Condition, wherein participants became familiar with the virtual environment and the task. If the number of errors in this phase was less than 8-10%, participants had to repeat the training and were requested to be faster. Participants performed 160 trials in total (96 incompatible, 64 compatible trials). Important, in CM Condition the movements of the avatar were in all cases the same as those executed by the participant which was expected to induce a strong illusion of ownership and agency with the avatar. At the end of this phase participants were required to complete a questionnaire in order to evaluate the experience of body ownership and agency towards the virtual body (see below).

After a short break the InCM Condition started. The InCM was divided in two blocks of 640 trials in each block (approximately 15 minutes each). During the InCM, in some infrequent trials participants' were deceived regarding their actions: that is, participants' correct movements to the visual target were falsified by an incorrect hand movement of the avatar, violating participant's internal intentions and causing (false)-alien-errors. The total amount of alien-errors matched approximately the percentage of natural self-generated errors in compatible trials (approx. 5-6% of trials, 64 trials in all participants). We avoided introducing (false)-alien-errors in the incompatible trials because these trials tend to pre-activate incorrect motor response tendencies and therefore the evaluation of the alien-errors could have been unclear for the participant (in some case, incorrect motor pre-activation might have leaded to an incorrect response that would have matched the incorrect alien-error) . To accomplish this effect of inserting alien-errors, we calculated the symmetrical position of the hand respect the yz plane, which is the correspondent to inverse the x value of the position. During the alien-errors participants did not notice anything until they started moving, since the avatar hand position was also at the initial position (0,0,0) during these trials (see an example in Fig S1, dotted grey line exemplifying the InCM condition leading to an alien-error). Correct responses, self-generated and (false)-alien-errors were all performed exactly at almost the same speed and timing (without noticeable delay). This task design allowed the comparison of behavioral and ERP effects related to three specific conditions: (i) correct responses towards the target; (ii) when the avatar was introducing (false) alien-errors; and (iii) those situations in which participant really performed an error (self-generated errors).

After two weeks some participants of the main experiment (n=9) returned to the lab to complete a control ERP experiment (Exp. 2). The Exp. 2 was carried out to inspect if the ERP effects related to (false)-alien-errors would be in tandem with to error-related brain processes associated with error observation (van Schie et al., 2004). In the control Exp. 2 the setup was the same as in the ERP Exp. 1: all participants were immersed in the virtual environment from a 1PP in the HMD. However in this occasion participants did not move and only observed the pre-recorded task performed during their own session in the Exp. 1. The pre-recorded movements were first cleaned from incomplete or corrected movements by substituting them for complete movements of the same participant, i.e. participants only visualized errors or correct responses. In order to avoid covert errors of the observer, the flanker arrows were removed and only the middle arrow remained in the screen (for similar procedure see van Schie et al., 2004). Participants were asked to pay attention to the task and count the number of times the avatar performed an erroneous movement. They were also told that at the end of the observation task they would be asked if the pre-recorded performance was theirs or from other participants. The aim of these instructions was to increase their level of attention. Since participants were not instructed to actively perform task it was expected that they would not develop a feeling of ownership towards the avatar (Banakou et al., 2013). In Exp. 2 we were specifically interested to inspect ERP differences between Observed Correct responses and Observed Error responses.

### The Experience Questionnaire

In order to evaluate the illusion of body ownership, localization and agency towards the virtual body participants were instructed to complete a 9-item questionnaire after each of the CM and InCM conditions. The questionnaire was adapted from previous studies (Botvinick & Cohen, 1998; Longo, Schuur, Kammers, Tsakiris, & Haggard, 2008) and contained a set of assertions that were scored with a 7-point Likert scale ranging from “strongly disagree” (-3) to “strongly agree” (+3). The questions were:

**Q1.** It felt as if the virtual body was my body (*myBody*)

**Q2.** I felt as if my hand was located where I saw the virtual hand to be (*collocated Hand*)

**Q3.** It seemed as if I might have had more than one body (*more Than One Body*).

**Q4.** It seemed as if the position of the hand I was feeling came from somewhere between my own hand and the virtual hand (*dislocated Hand*).

**Q5.** Most of the time, the movements of the virtual hand seemed to be my movements (*my Movements*).

**Q6.** Sometimes, I felt that the movements of the virtual hand were influencing my own .movements (*Influence*).

**Q7.** Sometimes, the virtual hand seemed to be moving by itself (*not My Movements*).

**Q8.** It sometimes felt as if my real hand was turning ‘virtual’ (*my Hand Virtual).*

**Q9.** It seems sometimes that the errors were not caused by myself (*not My Errors*).

Q1 and Q3 were related to the sense of body ownership. Increase in body ownership should be related to high scores in Q1 and low scores in Q3. Q2 and Q4 were related to the sense of proprioception and localization of the hand that participants experienced. Q5 and Q7 were related with visual-motor integration processes and violation of the sense of agency, important for the evaluation of the effectiveness of our experimental manipulation. We expected to record higher scores in Q5 and lower scores in Q7 in the CM condition when compared to the InCM condition. Q6 was an exploratory question on motor performance to assess how the visualization of the virtual movements influenced participants' real movements. Indirectly, this question assesses as well agency_violation. Q8 was a filler question about which we had no expectations. Q9 was a consistency check for the task performance in each condition, evaluating if participants were able to differentiate their own errors from the external errors. Q9 was expected to be higher in the InCM than the CM condition.

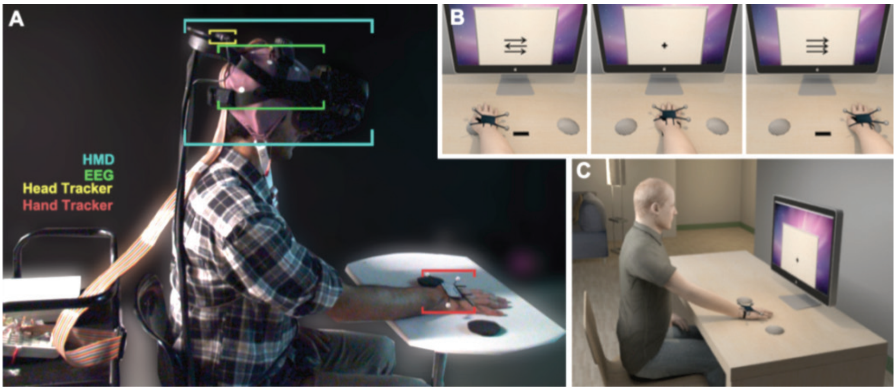


**Figure 1**. Experimental design used in Experiment 1. A. Participant in the laboratory with the head mounted display (HMD), electroencephalography (EEG) and the head and hand tracking systems. B. First Person Perspective (1PP) of the virtual arrow flanker task. Participants were instructed to perform fast movements with the right hand in the direction of the central arrow. After each movement the hand returned to the starting position (middle panel). The virtual hand followed the tracked real hand, but in some trials the displayed virtual hand movement was incongruent (InCM) with the participants' real movements, thus generating an alien-error. Three conditions were relevant for the EEG analysis, correct responses, self-generated (real) errors and alien-errors. C. Gender matched virtual body of the participant in the Immersive Virtual Environment (IVE).

### EEG Recording and ERP analysis

EEG was recorded from tin electrodes mounted in an elastic cap and located at 27 standard positions (Fz, F7/8, F3/4, Fc1/2 Fc5/6, Cz, C3/4, T7/8, Cp1/2, Cp5/6, Pz, P3/4, P7/P8, Po1/2, O1/2). All scalp electrodes were referenced offline to the mean activity of the left mastoid. Vertical eye movements (Electrooculogram, EOG) were monitored with electrodes located above and below the right eye. Horizontal EOG was collected from electrodes located at the outer canthus of each eye. Both vertical and horizontal EOG were used for artefact rejection and corrected with automated ICA-based methods. Impedances were kept below 5 kOhm. The electrophysiological signals were filtered with band-pass of 0.1–70 Hz (half-amplitude cutoffs) and digitized at a rate of 250 Hz. Trials with amplitude of more or less than 100 µV were rejected off-line.

ERPs time-locked to the onset of the response were averaged for epochs of -300 to 600 ms and with baseline set from -50 to 0 ms to the onset of the movement to the target. For the behavioural and ERP analysis only correct and error responses enter the analysis and error-correction movements were excluded. A low pass filter (14 Hz, half-amplitude cutoff) was applied in all computations. Trials following resting periods, self-generated and avatar-errors, and trials with RTs shorter than 150 ms or longer than 2.5 standard deviation of the individual RT mean were excluded from the analysis.

Mean ERP voltages were analysed by a three-way repeated measures ANOVA with factors Correctness (correct, self-generated error, avatar-error) x anterior-posterior electrode location (frontal, central and parietal) x lateral scalp location (left=3, midline=z, right=4). Time windows for statistical analyses of ERP voltages were chosen based on visual inspection of the grand average waveforms. We focus our attention on two distinct ERP waveforms arising after self-generated errors (at about 80-120 ms) and (false)-alien-errors (at about 310-360 ms).

For the ERP analysis regarding the control Exp.2 ERPs were locked to the onset of the observed responses performed by the self-represented avatar and averaged for epochs of -300 to 600 ms, with baseline set from -100 to 0 ms to the onset of the avatar responses. We were specifically interested to see differences related with Observed Correct responses and Observed Error responses. Since participants were unable to distinguish self-generated errors from avatar-errors, the bin of Observed Errors contained both type of errors. Mean ERP voltages were statistically analysed by a three-way repeated measures ANOVAs with factors Correctness (observed correct, observed error) x anterior-posterior electrode location (frontal, central and parietal) x lateral scalp location (left=3, midline=z, right=4) on the time window of 300-360 ms (time window where ERP effects related with alien-errors were observed).

## RESULTS

### Assessing illusions of ownership towards the self-represented avatar

Non-parametric Wilcoxon matched-pairs signed-rank tests were used to assess participants' questionnaire scores related to the experience of ownership towards the avatar during the Exp. 1 (Fig. 2, shows the medians and interquartile ranges). Regarding the Exp. 1 overall during the CM condition, when the avatar movements corresponded to those of the real body, a strong feeling of body ownership, localization and agency was induced. These results are consistent with previous findings (Banakou et al., 2013;Kalckert and Ehrsson, 2012;Sanchez-Vives et al., 2010). However, participant's sense of agency was impaired or partially disrupted when avatar-errors began to appear in the InCM condition as demonstrated by the scores in question, Q5, Q6, Q7 and Q9 (Fig. 2). The sense of body ownership, Q1 scores "It felt as if the virtual body was my body", was high in both CM and InCM conditions, with no significant difference between the two conditions ($Z = 1.519$, $p = 0.13$). However, the control question for body ownership (Q3) was significantly lower than Q1 in both CM ($Z = 3.682$, $p = 0.0002$) and InCM condition ($Z = 2.489$, $p = 0.013$); the global comparison between the Q1 and the control question Q3 was significant ($Z = -4.2857$, $P < 0.0001$). Interestingly, even in the InCM condition all participants still reported a strong and stable experience of body ownership. The occasional divergence between real and virtual hand locations during avatar-errors also had an effect on the hand localization item (Q2), the scores for the CM were greater than for the InCM condition ($Z = 2.902$, $p = 0.004$). Regarding visual-motor integration processes associated with participants' sense of agency, Q5 scores ("the movements of the virtual hand seemed to be my movements") were significantly higher in the CM than the InCM condition ($Z = -3.266$, $P = 0.0012$). In contrast the control question Q7 ("the virtual hand seemed to be moving by itself") scored higher in the InCM condition than the CM ($Z = 3.648$, $P = 0.0003$). In the InCM condition Q7 was marginally greater than Q5 ($Z = -1.963$, $P = 0.05$). This result shows that participant's sense of agency was impaired by the intromission of avatar-errors in the InCM. The InCM condition also influenced significantly more the participants' feeling about their movements (Q6) than the CM one ($Z=1.978$, $p = 0.048$). Finally, it is worth mentioning that in the InCM condition participants were aware that the errors introduced by the avatar were not their own errors (Q9) (for the comparison InCM vs. CM, $Z=3.543$, $p = 0.0004$).

During the control Exp. 2 (see supplementary material and Fig. S3) participants revealed a decreased experience of body ownership when Q1 scores were compared to the Exp. 1 CM condition's scores ($Z= -2.032$, $p = 0.042$), suggesting that the absence of movement (visual-motor asynchrony) was disrupting the experience of body ownership. Additionally, the proprioceptive consistency about the hand localization item (Q2) was also significantly diminished in Exp. 2 when compared to the Exp. 1 CM condition ($Z= -2.687$, $p < 0.\ 007$). Regarding visual-motor integration (Q5) the comparison shows a lost in the sense of agency during Exp. 2 when compared to Exp. 1 CM condition ($Z= -2.536$, $p < 0.\ 011$). Further analysis on (Q5 vs. Q7) showed significant differences ($Z= -2.570$, $p = 0.012$), meaning that overall, participants noticed that they could not control the avatar movements, thus showing a low sense of agency. These results are consistent with our expectations and show a low level of embodiment towards the self-represented avatar when compared to the Exp. 1 scores since the three most commonly described aspects of embodiment - body ownership, localization and agency - had low scores during the observation control Exp. 2.

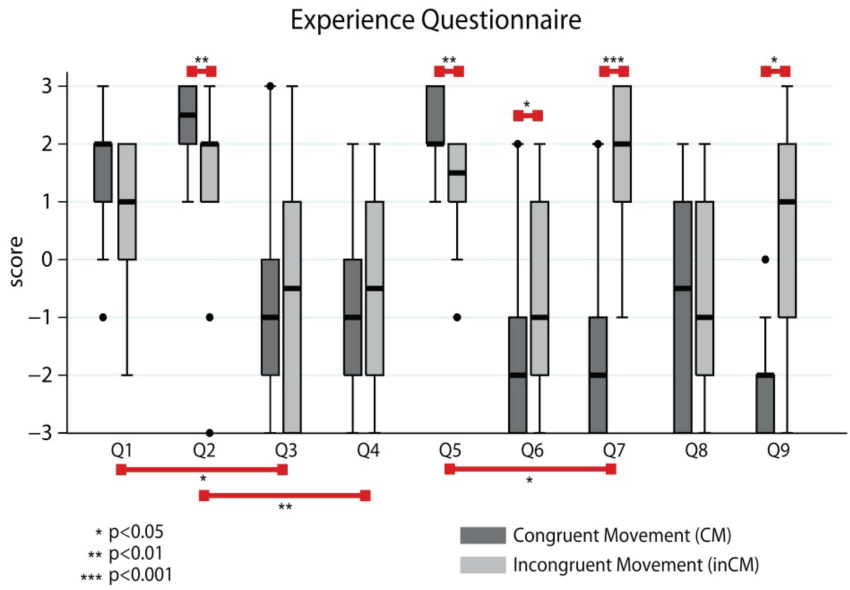


**Figure 2.** Results (mean ± SEM) of the questionnaire assessing the feeling of ownership and agency during CM and InCM conditions in Exp 1 [7-point Likert scale:"strongly disagree"(-3) to "strongly agree"(+3)]. Significant differences were tested with non parametric Wilcoxon signed-rank tests. **Q1**=*It felt as if the virtual body was my body*; **Q2**=*I felt as if my hand was located where I saw the virtual hand to be*; **Q3**=*It seemed as if I might have had more than one body*; **Q4**=*It seemed as if the position of the hand I was feeling came from somewhere between my own hand and the virtual hand*; **Q5**=*Most of the times, the movements of the virtual hand seemed to be my movements*; **Q6**=*Sometimes, I felt that the movements of the virtual hand were influencing my own movements*; **Q7**=*Sometimes, the virtual hand seemed to be moving by itself*; **Q8**= *It sometimes felt as if my real hand was turning 'virtual'*; **Q9**=*It seems sometimes that the errors were not caused by myself.*

## Behavioral performance on the virtual Flanker task

The mean percentages of self-generated errors produced by the participants were equal in the CM (mean + SD, 17 ± 9%) and InCM (17 ± 8%) conditions ($t(17) < 1$); thus the inclusion of alien-errors did not have a major impact on overall performance. As in other versions of the reaction time Erikson flanker task, our manipulation in the virtual environment revealed that participants were more accurate and faster responding to compatible trials compared to the incompatible ones: accuracy (compatible trials: 91.5 ± 6.9%; incompatible trials: 77.1 ± 10.5%, $t(17) = 11.302$ $p < 0.001$) and mean reaction time for correct responses (compatible: 259 ± 36 ms; incompatible trials: 273 ± 44 ms, $t(17) = -4.48$, $p < 0.001$).

We also investigated the extent to which compensatory cognitive control mechanisms, such as slowing down the reaction time in trials following errors (i.e. *post-error slowing*) (Logan & Crump, 2010; Marco-Pallares et al., 2008; Rabbit, 2002), were triggered after self-generated and alien-errors. In accordance with previous studies, self-generated errors compared to correct responses were followed by slower reaction times (e.g., 270 ± 46 ms vs. 262 ± 42 ms, respectively; $t\,(17) = 2.7$, $P = 0.027$). Strikingly, post-error slowing effects were larger after alien-errors (292 ± 51 ms; comparison with correct trials, 262 ± 42 ms: $t\,(17) = 7.1$, $P < 0.001$). The greater post-error slowing effect after alien-errors (~30 ms) compared to self-generated errors (8 ms) is surprising considering that during alien-errors participants were performing the correct movement. This result suggests a great impact in performance when participants detected a discrepancy between the expected (correct) output of their motor command and the observed (incorrect) movement performed by their virtual body self-representation.

### Neurophysiological mechanisms of error-monitoring in self-generated vs. alien-errors (Exp. 1)

ERP responses were compared in two specific conditions: when the virtual body was performing or introducing alien-errors and in those situations in which the participant really performed an error (self-generated errors) (see Fig. 3 and supplementary Fig. S2).

Self-generated errors when compared to correct responses gave rise to the classical frontal error-related negativity (ERN) component (see *blue line* in Fig. 3). The ERN signal peaked at about 100 ms after the onset of the erroneous movement and was maximal at frontocentral locations of the scalp. A repeated-measures ANOVA showed a main effect of Correctness [correct, self-generated error, alien-error)] ($F(2,34) = 12.9$, $P < 0.001$) as well as a significant interaction between Correctness per antero-posterior electrode location (frontal, central, parietal midline locations) ($F(4,68) = 6.6$, $P < 0.001$; showing a maximal effect of the ERN signal at frontal electrodes of the scalp for self-generated errors). This component is supposed to reflect the output of the internal–EMS related to error-monitoring processes

Surprisingly and in contrast to self-generated errors, alien-errors imposed by the alien agent elicited a negative ERP signal, which peaked at about 350-400 ms after the response, and was maximal at parietal electrode locations (ANOVA: Correctness, ($F(2,34) = 14.4$, $P < 0.001$; Correctness x Electrode (midline), $F(4,68) = 11.8$, $P < 0.001$; maximum effect at the parietal electrode) (see *green line*, Fig. 3). An additional statistical analysis considering the interaction between Condition [(self-generated errors minus correct responses) vs. (alien-error minus correct responses)] x the 27 electrodes locations revealed a significant interaction ($F(1,17)=5.3$, $P < 0.001$), which demonstrate the implication of distinct neural sources in the generation of both ERP components (Picton, Lins, & Scherg, 1995) associated, respectively, with self-generated errors (the early frontal ERN signal) and alien-errors (the later negative parietal ERP signal). These results suggest distinct neural computations and processing latencies regarding the evaluation of self-generated and alien-errors, and reveal an electrophysiological signature associated with the external- EMS during agency violations.

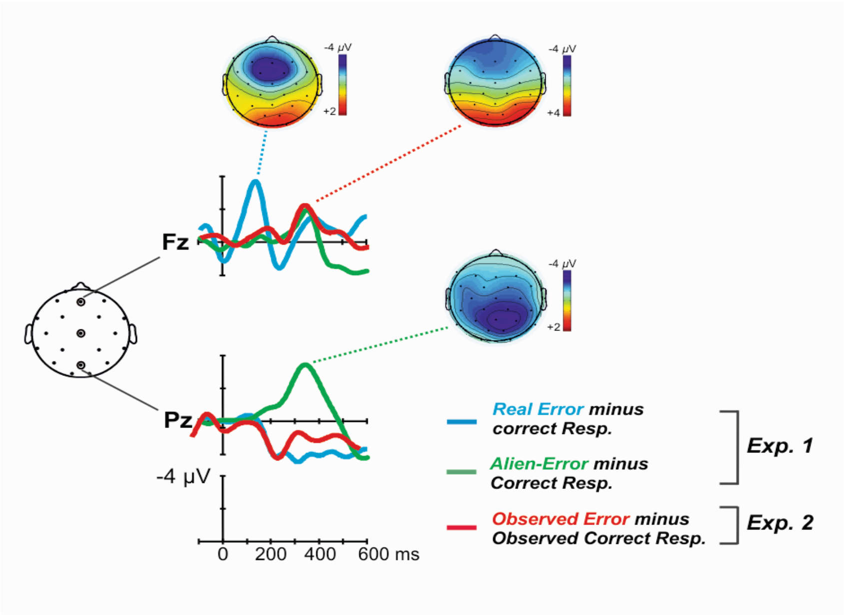


**Figure 3.** Main ERP results from Exp. 1 (participant executing the actions while embodied in a virtual body in a 1PP) and Exp. 2 (control experiment: participant observing the virtual body's actions in 1PP). Response-locked ERPs for the difference waveform computed subtracting self-generated errors vs. correct responses (blue line) showed the standard ERN. The ERN appeared immediately after the erroneous response (at about 100 ms) and showed a standard frontocentral topographical distribution. The comparison between alien-errors and correct responses (green line) yielded a completely different ERP component with a larger negative activity at 400 ms with a parietal topography. During the control ERP experiment (Exp. 2), the contrast of Observed Errors vs. Observed Correct Responses (red line) revealed a delayed ERN-like component at frontal electrodes.

Furthermore it was observed that the amplitude of this negative component over parietal regions (for the difference alien-error vs. correct responses) was negatively correlated with the subjective strength of virtual body ownership (computed as the difference Q1-Q3) ($r(18) = -0.6$, $P < 0.009$) (Fig. 4). Thus, the larger the subjective feeling of body ownership as measured by the subjective report, the strongest the amplitude of the negative parietal signal following agency violations.

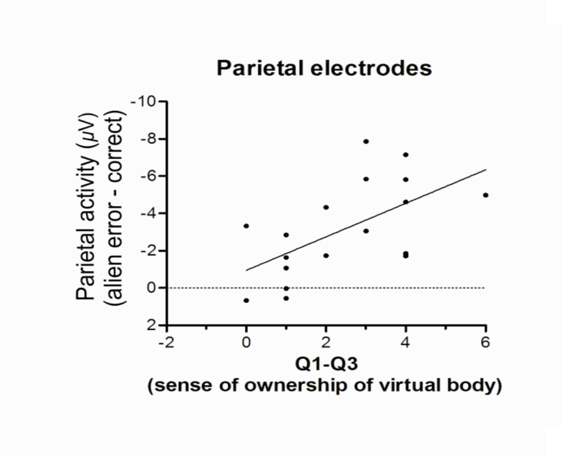


**Figure 4.** Correlation between the strength of the body ownership towards the virtual body and the negative parietal EEG activity related to agency violations ($r(18) = -0.6$, $P < 0.009$). The amplitude of this ERP signal was computed subtracting alien-errors minus correct responses in a selected parietal region including parietal electrodes (Pz, P3, P4 electrodes). The subjective strength of virtual body ownership was computed as the difference Q1-Q3.

### Neurophysiological signatures of error-observation in avatars (Exp. 2)

The ERP analysis of the control Exp. 2 revealed a delayed frontal ERN-like component when participants observed their self-represented avatar performing erroneous actions (Fig. 3, red line shows the contrast Observed error vs. Observed correct responses; see also supplementary Fig. S2), but no trace of the negative parietal signal associated with alien-errors observed in the Exp. 1. The delayed frontal ERN component developed at about 300 ms after occurrence of the observed erroneous action by the avatar and it was maximal at the frontal electrode (Fz) as indicated by a significant interaction Correctness (observed correct, observed error) x anterior-posterior electrode location ($F(2,16) = 26.4$, $p < 0.001$). The present results are in agreement with a previous study showing error-related brain activity in the medial-frontal cortex when participants have been exposed to errors performed by other agents (i.e. observational errors)(van Schie et al., 2004). This result rules out the possibility that the negative parietal signal elicited during avatar-errors in the active execution task (Exp. 1) could be due to mere error observation effects previously associated with the internal-EMS.

### DISCUSSION

Using ERPs we found that self-generated errors and not intended erroneous actions (false alien-errors), which affect one’s natural sense of agency, are associated with distinct neurophysiological signatures: showing different spatial cortical activity and different processing latencies. While, self-generated errors were followed by the frontocentral ERN signal, avatar-errors elicited a delayed negative signal in parietal regions of the brain. In a follow-up control ERP experiment it was shown that the neurophysiological processes underlying the monitoring of agency violations during alien-errors could not be attributed to monitoring processes related to error-observation. These findings suggest the existence of two independent, but not completely exclusive, error-monitoring systems in the human brain which are responsible for monitoring the correctness of self-generated and observed actions (internal-EMS) and in providing a coherent sense of agency over our actions (external-EMS).

Previous research in cognitive neuroscience has elucidated the existence of a highly flexible system in the medial-prefrontal cortex, which is involved in the rapid evaluation of ongoing events, being particularly sensitive to motor errors, negative outcomes and conflicting action-sets, and regulative aspects of cognitive and motor control (Carter, 1998;Danielmeier et al., 2011;Holroyd et al., 2005;Ridderinkhof et al., 2004). Our results are in agreement with this previous research by showing that the frontocentral ERN, a putative electrophysiological marker of the internal-EML, was modulated by the correctness of both self-generated (Exp. 1) and observed actions (Exp. 2) (van Schie et al., 2004).

More important, however, our results show that the internal-EMS is not responsible for processing the sense of control over one's actions, as the insertion of false alien-errors on participants' correct movements, which impaired their sense of agency and affected their normal performance (as reflected by greater post-error slowing effects), elicited a delayed negative ERP signal on the parietal cortex that was dissociated from the neurophysiological signature of the internal-EMS. Therefore, we disambiguate for the first time the neurophysiological processes underlying the monitoring of self-generated errors and errors which are caused without subjects' will.

It has been proposed that our ability to recognize ourselves as agents of our own behavior depends on constantly monitoring the sensory consequences of our actions. In normal everyday circumstances and running in the background of our consciousness, we experience an implicit and diffuse sense of coherence regarding the feeling of agency, mostly because there is a perfect congruence between the internal representations of our actions, the expected effects of the actions executed and the flow of resulting sensory events (multimodal reafferent feedback) (Farrer et al., 2008a;Pacherie, 2001). When a mismatch is detected between any of these internal and re-afferent signals a conscious feeling of agency violation might be triggered. Based on our findings we believe that the negative parietal ERP signal following (false) alien-errors reflect the output of this comparison process, which might lie at the core of the external-EMS. Furthermore, as the external EMS requires the processing of different re-afferent feedback information arriving at the somatosensory, visual and auditory regions with different neural delays, the comparison process might not finished until the degree of coherence is computed and a coherent multimodal representation is built. In line with our ERP data we believe that the timing required by the human brain to compute this comparison process may have a delay of 300-400 ms (~ time onset of the negative parietal ERP following alien-errors). Note that the error-monitoring process related to agency violations is slower than that the one computed by the internal-EMS (~ 100 ms) which may depend exclusively on central efference copy of the motor program independently of sensory or proprioceptive feedback information (Coles et al., 2001;Gehring et al., 1993;Rodriguez-Fornells et al., 2002).

Is important to mention that the parietal distribution of the negative ERP signal following alien-errors converges with the results of functional neuroimaging and lesion studies in which the importance of the inferior parietal cortex has been highlighted in relation to diminished feeling of agency (Farrer et al., 2003;Farrer et al., 2008b) and the comparison processes between intentions and action consequences (Desmurget et al., 2009;Sirigu et al., 2004). In line with these findings we propose that the external-EMS may be involved in higher-order aspects of motor control, such as awareness of action authorship (Farrer et al., 2008b), while the internal-EMS is involved in aspects related to goal attainment and adaptive modification of behavior (Ridderinkhof et al., 2004).

Interestingly, a significant association was observed between the amplitude of the parietal ERP signal and the subjective feeling of body ownership with the avatar. The larger the subjective feeling of body ownership as measured by the subjective report, the strongest the amplitude of the neurophysiological signature related to agency violations. This result suggests a relation between the internal semantic representations of the body image towards the self-represented avatar and the impact that unexpected imposed actions by the self-represented agent has on one's natural phenomenal experience of agency.

An intriguing question to be explored in future research is the exact computational nature of the parietal ERP signal following alien-errors. Interestingly this ERP resemble, in terms of latency and spatial distribution, the well-known N400 component, which has been classically associated to the violation of semantic and conceptual information (e.g., when listening to the sentence "I am going to eat a house") (Kutas & Federmeier, 2011). This component it has also been found to occur as a result of observing inappropriate motor plans (e.g., inserting screwdriver versus key into a keyhole) (Bach et al., 2009). Thus an interesting question that arises here is the extent to which the clash in the feeling of agency ("this is not my action") reflects a violation in the process of understanding our own actions (or our own "body action semantics"). In this sense, the comparison process underlying the ability to recognize ourselves as agents of our actions might not be too different from that carried out when comparing linguistic inputs or conceptual representations, as it might rely as well on the congruence of our own actions and the external consequences generated by these actions. Therefore observing the representation of an embodied body performing a non-planned action might be evaluated in a similar fashion as a semantic-conceptual violation. In a way, we are somehow observing "semantic violations of our body actions".

In sum, the present study provides crucial new evidence about how to distinguish at the neurophysiological level self-generated errors from errors which are caused without subjects' will and revealed, for the first time, an electrophysiological signature specifically related to agency violations. These findings shed new light on the timing required by the human brain to differentiate self-generated and externally caused errors, providing new neural evidences regarding the integration of internal and sensory feedback information in the build-up of a coherent sense of agency. Furthermore the new experimental setting presented here offers a unique opportunity to study the experience of agency while being embodied actors of our own actions, opening new avenues on current research focused on the mental processes underlying agency attribution in healthy and clinical populations.

.

## SUPPLEMENTARY MATERIALS:

**Movie S1.** This movie shows a participant in the laboratory fitted with the equipment (upper part of the video), entering the Immersive Virtual Environment (lower part of the video). Initially the participant explores the environment and experiences a new virtual body with visual-motor synchrony leveraged by the system head tracking and hand tracking. Then the experiment starts and the participant is asked to follow the direction of the center arrow of the flanker task by reaching with his hand a physical button. During the experiment the participant experiences some InCongruent trials. These trials are easily visible for an observer because the participant's hand and the avatar's hand diverge in opposite directions. This sequence of video contains a total of 40 trials: 38 are congruent (CM) and 2 InCongruent (InCM). InCongruent trials are mixed with Congruent trials and only occurred every 20 trials, approximately as infrequent than the natural errors produced during the main task in the compatible condition (5-6% of the trials). (see uploaded **video at** www.youtube.com/watch?v=lNjucylJxX8&feature=youtu.be).

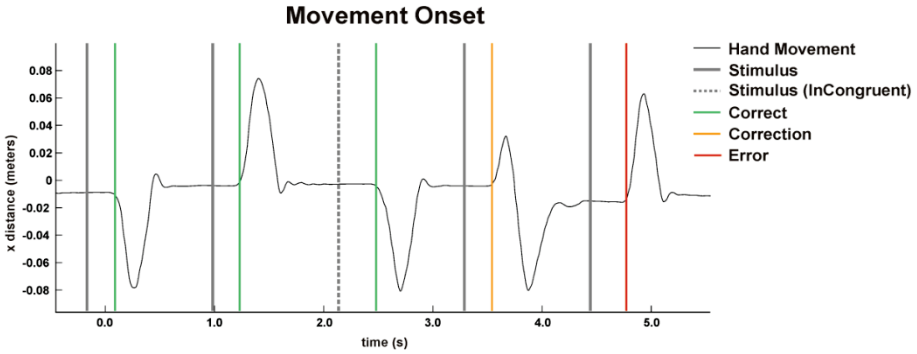


**Figure S1.** Section of the hand movements showing 5 trials; three of them are correct, one is a correction and one is a real error. Notice that one of the correct trials is an InCM trial (dotted grey line) to which the hand of the avatar went completely to the opposite direction, which result in a (false) alien-error.

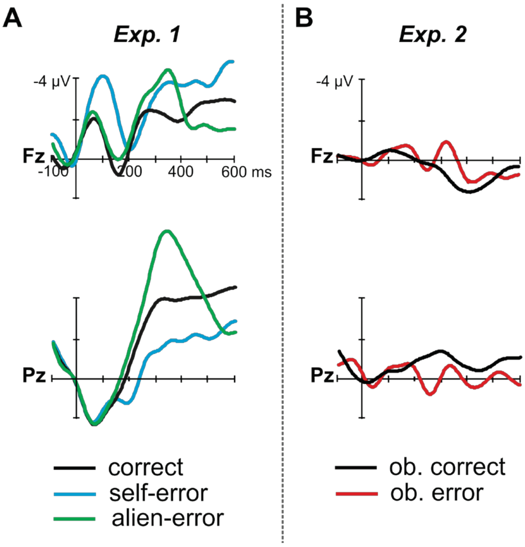


**Figure S2. A.** Response-locked ERP waveforms from main Exp. 1 associated with correct trials (black), real self-generated errors (blue) and false alien-errors (green). Depicted two electrode locations at frontal and parietal sides. **B.** Response-locked ERP waveforms from control Exp. 2 associated with observed correct trials (black) and observed error trials (red) at frontal and parietal sides.

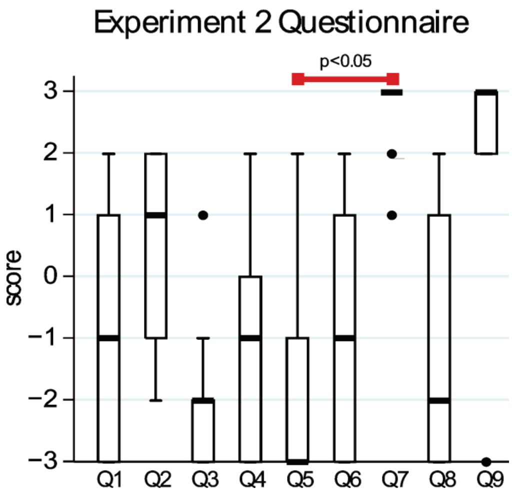
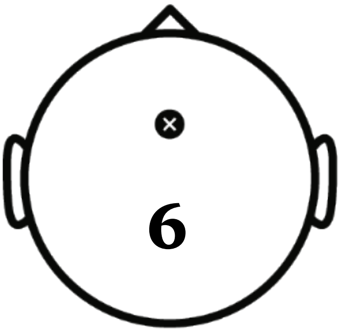


**Figure S3.** Score of the questionnaire after the observation task (Exp. 2)(*participant observing avatar actions in 1PP*). Global illusory embodiment with the virtual body was not found very intense, corroborated by the scores on the three relevant questionnaire items (Q1, Q2 and Q5) compared with control questions (Q3, Q4 and Q7) respectively (all *P* values > 0.05). The Boxplot shows the medians and interquartile ranges. Questions: **Q1** = *It felt as if the virtual body was my body*; **Q2** = *I felt as if my hand was located where I saw the virtual hand to be*; **Q3** = *It seemed as if I might have had more than one body*; **Q4** = *It seemed as if the position of the hand I was feeling came from somewhere between my own hand and the virtual hand*; **Q5** = *Most of the times, the* *movements of the virtual hand seemed to be my movements*; **Q6** = *Sometimes, I felt that the movements of the virtual hand were influencing my own movements*; **Q7** = *Sometimes, the virtual hand seemed to be moving by itself*; **Q8** = *It sometimes felt as if my real hand was turning 'virtual'*; **Q9** = *It seems sometimes that the errors were not caused by myself.*

**Experiment 2: Control ERP experiment** (participant observing avatar actions in 1PP)

Questionnaire scores

Non parametric Wilcoxon signed-rank tests were used to assess questionnaire scores (Figure S3). Participants (N=9) revealed a decreased experience of body ownership when Q1 scores were compared to the Experiment 1 Congruent scores (Z= -2.032, $p$ = 0.042), suggesting that the visual-motor asynchrony was disrupting the experience of body ownership. Additionally, the proprioceptive consistency about the hand localization (Q2) was also significantly diminished in Experiment 2 when compared to Experiment 1 (Z= -2.687, $p <$ 0. 007). Furthermore, when comparing the scores with the control question (Q2 vs. Q4) during all observations was diminished and not significant (Z= 1.807, $p <$ 0. 071) showing a divergence between the real and virtual hand localizations, as if the participant was having illusions of a dislocated limb. Regarding visual-motor integration (Q5) the comparison shows a lost in the agency during Experiment 2 when compared to Experiment 1 (Z= -2.536, $p <$ 0. 011). Further analysis on (Q5 vs. Q7) showed significant differences (Z= -2.570, $p$ = 0.012), meaning that overall, participants noticed that they could not control the virtual movements, thus showing a low sense of agency over the virtual hand. These results are consistent with our expectations and show a low level of embodiment in the virtual environment when compared to the Experiment 1 scores since the three most commonly described aspects of embodiment - body ownership, localization and agency - had low scores during the observation Experiment 2.



# 6

# Study 4

## Individual Differences in Error Tolerance in Humans: Neurophysiological Evidences [*]

---

[*] This study corresponds to:
Padrão, G., Mallorquí, A., Cucurell, D., Rodríguez-Fornells, A. Individual Differences in Error Tolerance in Humans: Neurophysiological Evidences. (**submitted**)

## INTRODUCTION

Humans have different tolerance levels when facing environmental errorful information. An important challenge for cognitive psychologists and neuroscientists is not only to understand the nature of the brain mechanisms and inherent cognitive processes underlying these differences, but also to explain how these differences are related to (in)adequate decisional processes carried out by human agents in their interaction with real-world environments.

The study of human error processing has advanced enormously during the last two decades, mostly due to the discovery of specific event-related brain components (ERPs) and neural networks that respond selectively to error-monitoring, error correction and compensation processes (Carter et al., 1998; Gehring et al., 1993; Holroyd & Coles, 2002; Ridderinkhof et al., 2004; for a review see Ullsperguer et al., 2014). However, most of this research has been focused exclusively on the examination of isolated performance errors (i.e., slips of action or lapses) in fairly simple reaction time tasks, whereby errors are carried out without conscious deliberation (Norman, 1981; Reason, 1990) and are mostly due to small attentional fluctuations or maladjustments in control (Botvinick et al., 2001; Gratton et al., 1988). However, there is a clear lack of neurophysiological research focused on the more complex error-prone scenarios or related to individual differences in on-line monitoring of complex chains of commands (Norman, 1988; Rasmussen, 1983; Reason, 1990; Woods, 1994). It is also worth noting that many catastrophes are often related to errors in human decision making processes during the operation of complex technological systems (e.g., drivers, air traffic controllers, etc.) (Casey, 2006; Wiegmann & Shappell, 2003; Woods, 1994). Thus it is important to move this research forward and to further investigate other factors that contribute to human failures in these settings (Reason, 1990). An attractive alternative is, for instance, to explore how individual differences in *error-tolerance* affects human decision processes and the extent to which these differences are explained by discrete neurophysiological mechanisms related to the monitoring and evaluation of error feedback in more uncertain and error-prone environments.

Error-tolerance makes reference to the capability of a system to continue functioning in accordance with design specifications even when undesired or unpredictable changes in the internal structure or external environment occur. At a cognitive level, in the presence of conflicting or error-prone information, error-tolerance could represent in humans a tendency or bias to follow well-programmed action plans, i.e. internal(rule)-based commands, absorbing new incoming information without the need to drastically change pre-selected courses of action. The concept of error-tolerance has been very important in the design of computer modeling of operating systems (Fields, Wright, & Harrison, 1995) and also in understanding the robustness of many complex systems that exhibit a great degree of tolerance against errors (Albert, Jeong, & Barabasi, 2000), including simple living organisms (Jeong, Tombor, Albert, Oltvai, & Barabasi, 2001). In human performance, error-tolerance has been investigated in low-level sensoriomotor learning, in which a large degree of noise, variability and redundancy exist in movement execution (Sternad, Abe, Hu, & Muller, 2011) as well as in speech motor control (Houde & Nagarajan, 2011).

In the context of decision-making, the study of error-tolerance might be particularly relevant in better characterizing the cognitive and motivational processes underlying the way human operators cope with errors and multiple feedback information in hazardous environments. For instance, individuals with a higher error-tolerance, whose actions tend to be mainly guided by internal(rule)-based strategies, might adopt strict patterns of behavior in more uncertain environments, following behavioral schemes that might have been successful in previous situations and showing an increased resistance to new errorful feedback information. These individuals might predominantly operate in a *feedback-independent fashion*. Moreover, due to their cognitive schemas individuals with higher or lower error-tolerance levels may commit different types of mistakes or diagnostic errors (Reason, 1990). An interesting type of higher-order errors in the study of error-tolerance in human performance are those described by Reason (1990; *see* also Rasmussen, 1983), the so called *rule-based mistakes*. Rule-based mistakes usually take place when the operator intentionally formulates a plan and carries it out using inappropriate rules (Norman, 1988; Reason, 1990). In this case, the agent might correctly identify the current goal but, through lack of knowledge, strength of the habit, effort or motivational disengagement applies an inappropriate rule.

An experimental paradigm particularly suited to assess error-tolerance in humans is the reversal learning (RL) task (see for example Jocham et al., 2009). In the RL task participants have to monitor different types of error feedback in order to learn stimulus-response associations that lead to rewarding outcomes. Critically, participants must adapt their behavior according to unpredicted changes in stimulus-response contingencies, being their decisions based on the trade-off of their rule-based knowledge (i.e., *the pattern of responses awarded up to a particular moment*) and the feedback received from the environment. In some cases this feedback can be randomly misleading, not rewarding the correct action or signaling the need for a behavioral change. Here, we examined the performance of 80 healthy participants during an RL task in order to select two extreme groups of individuals which showed a different propensity to maintain recently learned rules in spite of receiving different types of error feedback. Participants who perseverated with the newly learned rule for longer periods of time, independently of receiving error feedback, were assigned to the High Error-Tolerant (High ET) group. In contrast, those participants who exhibited more problems maintaining the newly learned rule after receiving negative feedback were assigned to the Low Error-Tolerant (Low ET) group.

To assess electrophysiological (ERP) differences in the impact of error feedback on performance between these two groups of ET participants we studied the activity of the feedback-related negativity (FRN) and the P3 component. The FRN has been associated with a performance monitoring system that evaluates the motivational significance of ongoing events, being especially sensitive to negative feedback and violations of expectancy (Gehring & Willoughby, 2002; Holroyd & Coles, 2002; Nieuwenhuis et al., 2004; Yeung, Holroyd, & Cohen, 2005). Modulations of the P3 component are associated with the updating of relevant (i.e. motivationally salient) external information in current task contexts involving the formation of adaptive decisions (Barcelo, Escera, Corral, & Perianez, 2006; Donchin & Coles, 1988; Fischer & Ullsperger, 2013), for example the internal accumulation of evidences that an error has occurred (Steinhauser & Yeung, 2010). Furthermore, we were interested in measuring differences in the degree of participant expectation or anticipation towards external feedback information using the stimulus preceding negativity (SPN). The SPN is an ERP component that is built up during the period preceding feedback presentation and has been described as an electrophysiological marker of subjects' anticipatory attention and motivational engagement to informative or relevant feedback (Brunia, Hackley, van Boxtel, Kotani, & Ohgami, 2011; Fuentemilla et al., 2013; Kotani et al., 2003; Moris, Luque, & Rodriguez-Fornells, 2013).

Based on the assumption that in uncertain environments, individuals with High ET, compared to Low ET, tend to favor internal-based commands - operating most of the time in a feedback-independent fashion and showing increased tolerance (resistance) to external feedbacks - we expected to observe a reduction in the motivational engagement and attentional impact during error feedback anticipation (indexed by a reduction on the amplitude of the SPN) and error feedback evaluation (reduction of the amplitude of the FRN and P3 components) in this group.

## METHODS

### Participants

Thirty-six right-handed participants were selected among 80 subjects from the University of Barcelona who underwent the ERP-reversal learning task. Participants were split in two groups (eighteen subjects each) based on their performance in the RL task (see below the categorization of High and Low Error-Tolerant groups). Both groups did not differ statistically in either age ($t(34) = 0.9$, $P > 0.05$) or gender ($\chi^2$ (1,36) = 0.64, $P > 0.05$). Participants reported no history of neurological or psychiatric disorder or drug consumption. All participants signed an informed consent and received monetary compensation that varied in accordance to their performance accuracy [23.3 ± 2.1 € (mean + SD)].

### Experimental design

A modified version of the probabilistic RL task used by Jocham et al. (2009) was used (*see* schematic description of the different conditions in Table 1). The task consisted of 1260 trials divided in 63 blocks with 16 to 24 trials each (pseudorandom order). In each trial, two grey squares located on either side of a central fixation point were presented over a black background for 1000 ms. Participants had to select one of the two squares by pressing one of two mouse buttons that corresponded to each stimulus. 700 ms after the response (*delay period-feedback anticipation*), a feedback stimulus indicating a win or loss of 0.06 € (*feedback processing*: happy or sad face respectively) was presented in the middle of the screen for 800 ms (*see* Fig. 1a). Inter-trial stimulus was set to 500-900 ms (randomized) to avoid automatic responses. When participants did not respond within 1000 ms after the presentation of the two squares a question mark appeared in the screen. Self-paced resting breaks for free blinking were given every 3 blocks coinciding with the 7th-10th trial of each block to avoid any participants' anticipation of the rule change. Information about the amount of money gained until that moment was provided during these breaks.

Participants were informed of the existence of two probabilities of winning and losing (75% and 25%) associated with each stimulus. On each trial participants had to choose the stimulus leading to the reward outcome. After a randomly jittered block length of 16–24 trials reward contingencies were reversed (rule reversal) and participants had to switch their selection to the new rewarded alternative. Thus, due to the 75% probability of winning, participants had to monitor different types of negative (error) feedback, which in some cases could be misleading, non-rewarding feedback of a correct response or signaling the need of a behavioral change (when concurring with the rule-reversal). Crucially, participants were instructed to change the response pattern only when they were completely sure that the rule had changed. A brief training session was conducted at the beginning of the session to ensure comprehension of the task. Participants were encouraged to increase their earnings as much as possible during the task.

### Selection of the participants based on Error Tolerance

From the 80 participants that performed the RL task we selected two extreme groups (each composed by 18 participants) based on error-tolerance. Error-tolerance was defined as the ability of a participant to continue following the current rule after a successful reversal despite the presence of misleading or spurious error feedback (note that only 75% of correct trials were rewarded).

To select the members of each of the Error Tolerant (ET) groups we inspected the behavioural performance of all 80 participants throughout the eleven consecutive trials following a successful behavioural switch (i.e. when the rule has actually reversed and participants had to switch to the new correct response) (*see* Fig 1b);

Specifically, we plotted the data of all 80 participants showing the proportion of trials after reversal in which the new learned rule was maintained throughout the eleven consecutive trials (averaged across all 62 experimental blocks). The most extreme participants (n=18 per group) who persevered with the new learned rule for longer periods of time, independently of the presence of spurious (error) feedback or error feedback informing of a rule-change, were assigned to the High ET group. In contrast, the 18 participants that exhibited more problems in maintaining the new learned rule, swinging between the response alternatives more often after receiving error negative feedback, were assigned to the Low ET group (*see* Fig. 1b).

### EEG recording

EEG activity was recorded using tin electrodes located at 29 standard positions positions (Fp1/2, Fz, F7/8, F3/4, Fc1/2 Fc5/6, Cz, C3/4, T3/4, Cp1/2, Cp5/6, Pz, P3/4, T5/6, PO1/2, O1/2) mounted in an elastic cap. Vertical eye movements were monitored with an electrode at the infraorbital ridge of the right eye. Electrode impedances were kept below 5 kΩ. The electrophysiological signals were digitized at a rate of 250 Hz and filtered with an online band-pass of 0.01–70 Hz (half-amplitude cut-offs). All scalp electrodes were referenced offline to the mean activity of the left and right mastoids.

### Behavioural analysis

Different types of events were defined for the behavioural (and ERP) analysis of the RL task (see Table 1 for clarification):

i. <u>Spurious negative feedback</u> (**SN**). An unexpected (misleading) negative (error) feedback (monetary loss) received by the participant after a correct rule response, due to the 25% probability of receiving an error or incorrect feedback. Importantly, we included only those trials that were preceded by at least 5 positive feedback trials, which indicate that the current rule was correctly consolidated for that particular block.

ii. <u>Reversal trial</u> (**RT**). The *first negative error feedback* concurring with a rule shift and a new starting block.

iii. <u>Final reversal error</u> (**FRE**). The *last negative error feedback* after which the participant switched his/her response pattern to the correct new rule.

iv. <u>Positive feedback</u> (rewarded) trials. Rewarded responses after a correct response which followed the expected rule. For the ERP analysis only those trials that followed the 5th trial after the rule-change were selected to ensure that the new rule has been correctly consolidated.

Regarding the behavioural performance we also inspected the following events:

v. Reversal errors. Perseverative errors after a reversal trial. These errors (see table 1) occurred when the participant did not change the response pattern after the rule changed (Reversal Trial).

vi. Loss-shift errors. After receiving an error feedback which was not consistent with a rule shift (Spurious negative feedback), the participant changed the response incurring a monetary loss.

### ERP analysis

EEG data was low-pass filtered offline at 12 Hz for ERP analysis. ERPs related to anticipation/expectation toward the outcome (SPN) were studied by extracting response-locked EEG epochs of 900 ms starting 100 ms before the button press (baseline). ERP components associated with the feedback/outcome processing (FRN and P3) were investigated by epoching EEG data from -100 ms (baseline) to 600 ms, time-locked to the feedback onset. Trials from the First block were not entered into the analysis. Trials exceeding ± 100 μV in both EEG and EOG were excluded from further analysis.

The time windows for the statistical analyses of ERPs were chosen based on visual inspection of the grand average waveforms. The SPN was computed in the 200 ms period prior to feedback presentation (Moris et al., 2013). The FRN amplitude was averaged in a time window located 50 ms around the peak (290-340 ms). Because the FRN was superimposed onto a slow positive waveform (the P3), ERP epochs were further filtered using a band-pass filter (3-12 Hz) in order to remove the slow positive waveform in which the FRN develops and to stabilize possible drifts. The analysis of the P3 component was performed on the time window of peak activity, 350-500 ms.

Statistical effects regarding neural processes related to anticipation and evaluation of positive feedback (rewarded trials), co-occurring with correct rule-based responses, were also tested using repeated measures ANOVAs using factors of electrode location (frontral, Fz; central, Cz; parietal, Pz) and Group (Low ET and High ET) on the mean amplitude of the SPN (anticipation) and P3 (evaluation). Sequences of positive feedback trials were selected to inspect participants' tolerance to external cues that signal the expected, current (rewarded) rule.

Anticipatory processes (SPN) towards potential error feedback information were examined with ANOVAs with two within-subjects factors: error feedback type (Spurious negative, Reversal trial, Final reversal error) and electrode location (Fz, Cz, Pz), and one between-subject factor Group (Low ET and High ET). The analysis of different types of negative feedback was important in assessing differences in error tolerance in neural expectation processes towards more irrelevant or redundant

(Spurious negative events) and relevant learning error feedback (Reversal trial and Final reversal error).

The same method was applied to test statistical effects on the mean amplitude of the FRN and P3 components during error feedback evaluation. These two neural responses directly indicated the motivational impact of error feedback processing in different error tolerant systems.

For all statistical effects involving two or more degrees of freedom in the numerator, the Greenhouse-Geisser epsilon was used to correct possible violations of the sphericity assumption. P-value after the correction is reported.

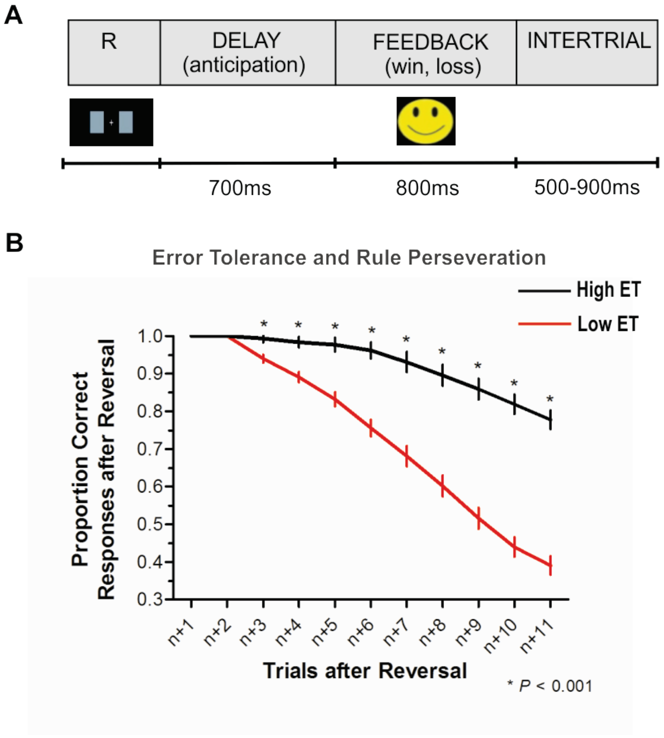


**Figure 1. (A)** Schematic illustration of the task design. For each trial participants had to choose one of the squares placed on each side of the screen by pressing the corresponding mouse button. After a delay period of 700 ms feedback indicating a monetary win or loss was presented. **(B)** Perseveration of the rule in High (black) and Low (red) ET groups. The X-axis shows the number of trials after a successful reversal of behaviour, i.e. immediately following the Final reversal error. The values on the Y-axis are the percentage across 11 trials (mean+SEM) in which the participants maintained the current rule.

## RESULTS

### Behavioural performance

Participants were divided into two groups based on their propensity to sustain the new rule after task contingencies changed and independently of the presence of spurious events (i.e. a distinctive feature of High ET group). To examine the differences in the behavioural patterns of both groups we first analysed the average of participants' correct responses (across all 62 blocks) throughout 11 trials ($n+11$) after the rule change. This measure gave us an index of the proportion of trials in which participants maintained the current rule across time. A repeated measure ANOVA was carried out including Trial number (from 1 to 11 after correct rule change) and Group (High ET and Low ET). A main effect of Trial number showed a clear overall decrease in the tendency to maintain the new rule across time ($F(10,340) = 233.9$, $P < 0.0001$) (Fig. 1b). The significant interaction Trial number x Group ($F(10,340) = 58.4$, $P < 0.0001$) confirmed that while the Low ET group exhibited drastic problems in maintaining the newly learned rule, the High ET group showed a greater tolerance to error feedback and therefore 'stuck to' the former rule for a higher number of trials. The Low ET group began to significantly differ from the High ET group from the third trial after the reversal (parwise comparisons *t-test* at all-time points, $P < 0 .001$).

The Low ET group showed a higher amount of Loss-shift errors [Low ET: 36 ± 23, High ET: 5 ± 4: $t(34) = 5.6$, $P < 0.001$]. These mistakes demonstrate the tendency of this group to change their response pattern right after the presentation of error feedback, even when the rule had not changed. In contrast, the proportion of Reversal errors was greater in the High ET group (High ET: 140 ± 18; Low ET: 96 ± 14: $t(34) = 8.2$, $P < 0.001$). On average, the High ET group needed more error feedbacks to change the current rule after a Reversal trial (High ET: 3.2 ± 0.35; Low ET: 2.3 ± 0.27: $t(34) = 8.2$, $P < 0.001$). This result clearly showed that High ET participants perseverated more on the previous rule during the consecutive trials (even when the current rule was no longer valid), taking more time to adapt their behaviour to the new task contingencies. In summary, the present results clearly illustrate that High ET participants exhibited less erratic response changes after error feedback, being less influenced by misleading error feedback presented during the task. However, due to their rule-maintenance bias High ET participants tended to perseverate on the activated rule for an extended period of time when compared to Low ET participants and were slower to change behaviour after rule reversal.

Interestingly, despite the discrepancy of the policies adopted by the two groups of participants, the amount of money collected along the task did not differ between them (High ET: 23.1 ± 2.4 €; Low ET: 23.1 ± 1.9 €, $P > 0. 5$).

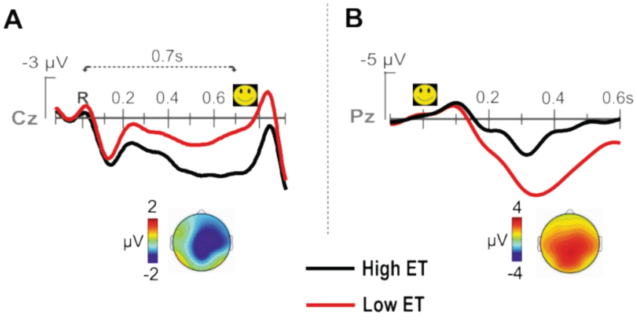


**Figure 2.** ERPs regarding **(A)** anticipation (SPN modulation at electrode Cz) and **(B)** evaluation (P3 modulation at electrode Pz) of positive feedback (rewards). Note the increase of the SPN and P3 components on the Low ET group. Also depicted are the topographical maps of the SPN (**A**, time window of 200 ms preceding feedback presentation) and P3 components (**B,** time window of 350-500 ms) for the difference between Low ET minus High ET.

## ERP Results

### *Anticipation and processing of expected Positive Feedback*

To evaluate differences in error-tolerance related to the anticipation and processing of expected positive outcomes that signalled the current (rewarded) rule selection we investigated the amplitude of the SPN and P3 components. Statistical effects on the SPN mean amplitude (200 ms period preceding feedback presentation) were evaluated with an ANOVA including Electrode location (Fz, Cz, Pz) and Group (Low ET and High ET) (Fig. 2a). A main effect of Group was encountered ($F(1,34) = 7.5$, $P < 0.05$) showing a decreased amplitude of the SPN in the High ET group (Fig. 2a). This group difference was maximal at location Cz (Electrode location x Group interaction, $F(1,34) = 6.6$, $P < 0.05$; *see* the scalp distribution of the SPN component for the difference Low ET-High ET, Fig. 2a). This result suggests that the amount of expectation and motivational engagement towards the information conveyed by the imminent feedback was lower in the High ET group.

Complementing the previous results, a clear decrease was observed in the P3 amplitude (350-500 ms) for the High ET group (main effect of Group, $F(1,34) = 13.8$, $P < 0.001$; see Fig. 2b) after the presentation of the positive feedback. The differences between both groups were larger at Pz (electrode location x Group interaction, $F(2,68) = 3.9$, $P < 0.05$; see the centro-parietal distribution of the P3 in Fig. 2b). This reduction in the amplitude of the P3 was expected considering the reduced motivational impact that expected feedback might have on participants' whose decisions are internal-rule based (i.e, High ET group).

Importantly, a correlation analysis (Pearson coefficient) between the amplitude of the SPN and the amplitude of the P3 showed the expected relationship between both components. A larger P3 amplitude was observed for higher negativity in the

SPN (r = - 0.6, $P < 0.001$, computed at Pz location). This correlation is in agreement with the idea that the larger the expectancy towards external feedback, the larger the motivational impact of its presentation (Moris et al., 2013; Fuentemilla et al., 2013).

### *Anticipatory processes associated to error feedbacks*

To inspect differences in the motivational engagement of the participants before receiving relevant or potentially misleading error feedback, we first examined the modulation of the SPN component in response to different types of error feedback during reversal. An ANOVA analysis was conducted with factors error feedback type (Spurious negative, Reversal trial, Final reversal error), electrode location (Fz, Cz, Pz) and Group (Low ET and High ET). The High ET group exhibited an overall reduction of the SPN compared to the Low ET group (main effect of Group, $F(1,34) = 4.9$, $P < 0.05$) (see Fig. 3a and 4a). The differences between groups were higher at Cz (Electrode location x Group interaction, $F(1,34) = 9.4$, $P < 0.005$; see in Fig. 4a, the scalp distribution of the SPN for the difference between both groups).

Importantly, we observed that the SPN amplitude was modulated by the relevance of the feedback (main effect of error Feedback type, $F(2,68) = 28.5$, $P < 0.001$), being greater (i.e. more negative) just before Final reversal errors (FRE), that is the last error feedback preceding the behavioural shift (Fig. 3a and 4a). The increase of the SPN during Final reversal errors is important because it signals the moment of updating of the current rule and, naturally, the period in which participants may generate greater expectancies about the incoming result. A marginal interaction of error feedback type x Group ($F(2,68) = 2.9$, $P = 0.065$) showed that while the anticipation of spurious errors (i.e. Spurious negative and Reversal trials) was different in the participants' error-tolerance, the motivational engagement just before relevant error feedback leading to behavioural switching (Final reversal errors) increased in both groups equally (see Fig. 4a and the diminished difference between groups in the corresponding topographical map of the SPN). This result was confirmed with further pairwise t-tests comparisons (at location Cz). Significant differences in the SPN amplitude between both groups were observed during Spurious negative ($t(34)= 3.15$, $P < 0.005$) and Reversal trials ($t(34)= 3.5$, $P < 0.005$), but no differences were observed during Final reversal errors (the SPN amplitude being similar in both groups, $t(34) = 1.1$, $P > 0.05$).

### *Processing of error feedback*

The analysis of the processing (evaluation period) of error feedback showed a very similar pattern to the one observed for the SPN. A similar ANOVA design was used:

error feedback type (Spurious negative, Reversal trial, Final reversal error) x electrode location (Fz, Cz, Pz) x Group (Low ET and High ET).

Overall, the amplitude of the FRN was lower in the High ET group when compared to the Low ET group (main effect of Group: $F(1,34) = 9.5$, $P < 0.005$), confirming our predictions that subjects with High ET are in general less responsive to error feedback (Fig 3b and 4b). The FRN was maximal at electrode Fz (main effect of electrode, $F(2,68) = 35.8$, $P < 0.0001$; see the classic fronto-central distribution of the FRN at Fig. 4b) and its amplitude was modulated by the type of error feedback ($F(2,68) = 17.5$, $P < 0.0001$): the FRN was greater (i.e. more negative) during more unexpected feedbacks (Spurious negative and Reversal Trials) and reduced in response to the most expected ones, i.e. Final reversal error (see the bottom of Fig. 3b and Fig. 4b). Interestingly, no group differences were found during the evaluation of Final reversal errors as exposed by a significant error feedback x Group interaction ($F(2,68) = 4.9$, $P < 0.05$). Further t-tests comparisons (at Fz location) showed that group differences were present only during Spurious negative ($t(34) = 3.2$, $P < 0.005$) and Reversal trials ($t(34) = 3$, $P < 0.01$) but disappeared during the Final reversal errors ($t(34) = 1.4$, $P > 0.05$) (Fig. 4b). The absence of group differences during final reversal errors suggests that the level of expectancy towards the last error feedback before correct rule-based behavioural adjustments was identical in the two groups.

In line with the FRN results, it was seen that the P3 amplitude was overall lower in the High ET (Group effect, $F(1,34) = 8.2$, $P < 0.05$). Importantly, the difference between groups decreased during Final reversal errors as suggested by the interaction error feedback type x Group interaction ($F(2,68) = 4.6$, $P < 0.05$) (Fig. 3b and 4c). This result suggests that Final reversal errors were similarly salient in both groups. It is important to mention that a main effect of error feedback type was observed for the P3 ($F(2,68)=22.7$, $P < 0.0001$), indicating that larger P3 amplitudes were developed after error feedbacks that were directly followed by a behavioural switch (Final reversal errors). Further t-test comparisons at location Pz, which was the electrode with the highest P3 amplitude ($F(2,68) = 35.6$, $P < 0.0001$), revealed that the P3 amplitude did not differ significantly between groups during the Final reversal errors ($t(34) = 1.6$, $P > 0.05$) but statistically differed during Spurious negative ($t(34) = 3.4$, $P < 0.005$) and Reversal Trials ($t(34) = 2.9$, $P < 0.05$; see also Fig. 4c, for the centro-parietal distribution of the P3 component for the difference between groups).

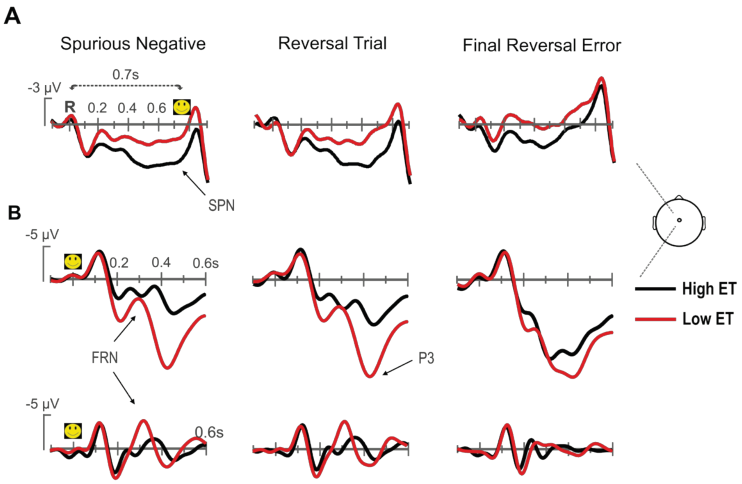


**Figure 3.** ERPs at electrode Cz during the anticipation and evaluation of negative feedback in High (black) and Low (red) ET groups. **(A)** Development of the SPN component prior to feedback presentation (feedback anticipation). **(B)** Grand-average waveforms (FRN and P3) associated with negative (error) feedback evaluation: on the top the ERPs filtered with a low pass filter; on the bottom ERPs filtered with a 3-12 Hz band-pass filter in order to isolate the FRN component.

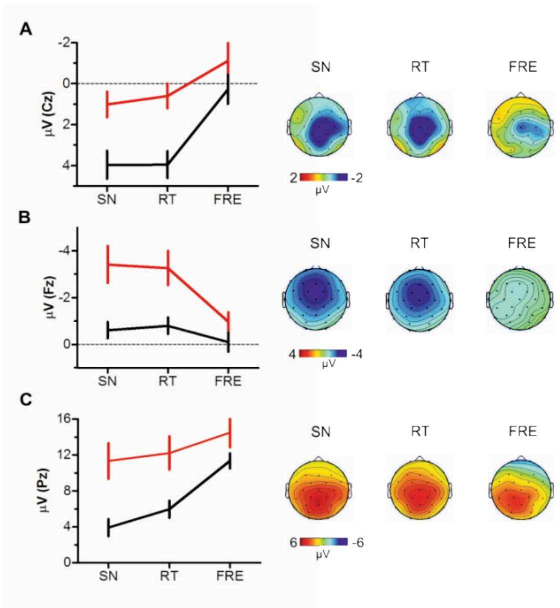


**Figure 4. (A)** Mean amplitude of the SPN in High (black) and Low (red) ET groups during the anticipation of negative error feedback (SN, Spurius negative; RT, Reversal trial; FRE, Final Reversal Error) at electrode Cz and topography of the SPN (average signal over the 200 ms prior to feedback presentation; difference Low ET - High ET). **(B)** Mean amplitude of the FRN (peak activity, 290-340 ms) at electrode Fz and the related topographical map (difference Low ET - High ET) during the evaluation of SN, RT and FRE. **(C)** Mean amplitude of the P3 (peak activity, 350-500 ms) at electrode Pz and the topographical maps illustrating the difference between groups (Low ET - High ET) during the evaluation of SN, RT and FRE . Error bars represent SEM.

## DISCUSSION

In the present study we investigated whether individual differences in error-tolerance were predictive of discrete neurophysiological mechanisms reflecting specific cognitive and motivational processes underlying the monitoring of error feedback presented in dynamic environments involving complex rule-base decisions. Using a reversal learning task we compared the neurophysiological (ERPs) responses of two extreme groups of participants with High and Low ET. This comparison showed differences in their propensity to perseverate (or change) on learned response patterns (i.e. rule-based learning) in spite of the presence of error feedback information, which could be either misleading, conflicting with on-going correct conduct or implying the reprograming of new rules and adaptive changes in behavior. It is worth mentioning that all participants were explicitly instructed to change their behaviour only when they were certain that a received error feedback was concomitant with rule-changes.

Despite the given instructions we found that one group of participants was more prone to maintaining current learned rules for longer periods of time, displaying a higher tolerance to misleading error feedback received during each learning block (see Fig 1b; block sequence example in Table 1).When these individuals, with High ET, “stuck to” a particular plan of action, they were more averse to changing their behaviour, even when error feedback was informing them that the current rule was no longer valid. As we predicted the High ET group showed a response pattern that could be characterized as feedback-independent in which the trial-by-trial feedback information has a lesser influence on subsequent decisions. Because of this, High ET was associated with an increased incidence of perseverative errors (i.e. reversal errors). According to Rasmussen & Vicent (1989), perseverative errors echo the application of inappropriate expectations or inefficient monitoring of environmental cues, probably due to the strength of the habit or motivational disengagement. The prevalence of perseverative errors on High ET systems most likely reflects a cognitive bias (*cognitive conservatorism or change-aversion*) to favour internal-rule-based commands rather than external (environmental) guided information, devaluing new incoming information and exploiting behaviors that have been successful in previous situations (e.g. well learned routines). In contrast, we observed that participants with Low ET exhibited more problems in maintaining a stable behaviour following the new learned rules. They were more willing to explore new alternatives in their environment after facing negative evaluations, consequently incurring inappropriate Lost-shift errors (i.e. mistakes caused by inadequate monitoring of spurious error feedback).

Interestingly, although High and Low ET groups showed different decision patterns in coping with errorful information, both ended the task with equal gains. This interesting result not only puts forward for consideration the existence of different thresholds of error-tolerance in humans but it also suggests that different decision making strategies can lead to similar final outcomes. In this sense, error-tolerance seems to be an important construct in understanding trial-by-trial variations in decision making patterns and to explain individual differences in the propensity to make mistakes in hazardous and fast-changing environments. To better illustrate this relationship between error-tolerance and (in)adequate decisional processes in complex real life environments, let us consider the situation of an operator with High ET in a flight simulator. Here they must handle and process multiple auditory and visual alarms and displays informing them of the current functioning and status of the machine or of potential risks and issues requiring solutions. Based on our behavioral data, we suggest that High ET is related with more conservative and tenacious decisions in environments in which multiple error feedbacks need to be constantly monitored. High ET individuals seem more likely to follow instructions and might be more "*protected*" against errorful scenarios or conflicting events (such as erroneous feedback). In this sense, High ET participants could withstand more negative feedback without getting distracted or changing pre-activated courses of action. This cognitive bias might predict a small prevalence of inadequate behavioral changes following erroneous or spurious feedbacks. However, due to their tendency to behave based on internalized rules, they may lose cognitive flexibility in certain circumstances - developing a sort of "*cognitive tunel vision*" (i.e. selecting a particular hypothesis and sticking to it without looking for alternatives) (Senders & Moray, 1991). In this particular case, High ET participants would fail to recognize important environmental cues demanding the implementation of new behaviors, resulting in them being less able to react to potential "threats" (Cañas, Quesada, Antoli, & Fajardo, 2003; Norman & Shallice, 1986; Rougier, Noelle, Braver, Cohen, & O'Relly, 2005; Spiro, Coulson, Feltovich, & Andersib, 1988). In this sense, although individuals with a higher error-tolerance may be more effective in monitoring potential misleading information, they seem to be more susceptible to missing sudden and important changes, as demonstrated by the increased number of perseverative errors on High ET participants.

Looking at the neurophysiological mechanisms involved in error feedback monitoring in different ET systems, it was shown that participants with High ET when compared to those with Low ET exhibited reduced SPN (Fig. 3a), FRN and P3 (Fig. 3b) activity during the anticipation and processing of error feedback that was not directly associated to rule-based behavioral changes (i.e. Spurious negative and Reversal trials). Interestingly, High ET participants also showed reduced amplitude of the SPN and P3 components on positive feedback that reinforced the current rule application (Fig. 2). The diminished SPN amplitude in the High ET group suggests less attention orientation and motivational engagement towards the information conveyed by external events that did not imply rule-based changes in behavior (Brunia et al., 2011; Kotani et al., 2003; Moris et al., 2013). Furthermore, the reduced activity of the FRN and P3 components in the High ET group shows that this group may perceive error feedback as motivationally less salient or informative for the effectiveness of on-going performance (Fischer & Ullsperger, 2013; Gehring & Willoughby, 2002; Nieuwenhuis et al., 2004; Picton, 1992; Polich, 2007; Yeung et al., 2005). These results agree with our proposal suggesting that individuals whose performance relies mostly on internal-rule based knowledge (i.e. High ET) are less responsive to and less dependent on external environmental information, operating predominately in a feedback-independent mode.

It is important to mention that the observed differences in EEG activity between the two groups cannot be entirely explained by group differences on the general processing of external feedback information. Critically, we found no group differences during the anticipation (SPN component) and evaluation (FRN and P3 components) of Final reversal errors (i.e. error feedback that was followed by correct rule-based behavioral switches) (Fig. 3 and 4). This result is very interesting as it shows that despite the fact that individuals with different degrees of ET are affected differently by potentially misleading or redundant error feedback information, that once they have collected enough internal evidence that task contingencies have actually changed, their neural patterns of EEG activity predicting behavioral changes are similar.

The observed SPN increase over centro-parietal regions prior to Final reversal errors (Fig 4a) is in line with previous findings showing an augmented amplitude of this component prior to the presentation of motivationally relevant information (Brunia, 2011; Fuentemilla et al., 2013; Kotani et al, 2003). However, the absence of group differences during this period demonstrates that regardless of individual differences in error-tolerance, error feedback is perceived as more salient to behavioral changes and tends to involve enhanced attentional and motivationally appetitive processes (Fuentemilla et al., 2013). Furthermore, similar neural computations of expectancy towards the last error feedback predicting behavioural-switches were also seen in both groups. The FRN amplitude was equally decreased in the two groups during final reversal errors (Fig 4b). Another important result was that although both groups accumulated a different number of error feedbacks after Reversal trials before deciding to switch to the new correct response alternative, the neural mechanisms leading to the subjective awareness of error commission prior to behavior adaption was marked by an equal increase of centro-parietal P3 activity in both High and Low ET groups. The increase of the centro-parietal P3 component during final reversal errors (Fig 4c) is in agreement with recent works in decision making suggesting that this neural signal is involved in the accumulation of internal evidences leading to the awareness/detection of relevant events requiring the formation of new behaviors (Fischer & Ullsperger, 2013; O'Connell et al., 2012). Behaviors such as the updating process leading to the awareness that an error has been made (Steinhauser & Yeung, 2010) and behavioral changes are required.

Together, these results suggest that even though individuals with High and Low ET display different degrees of attention, motivation and generate different expectations towards redundant external error feedbacks, they share common adaptive cortical mechanisms of updating error values when becoming aware that erratic trends in action require new behavioral solutions.

These neurophysiological evidences support discrete cognitive and motivational processes underlying human differences in ET and further suggest a close relationship between error-tolerance and different loops of cognitive control. In this regard, while High ET systems expose a strong cognitive bias to act in a feedback independent mode, favoring internal-based commands and showing an increased susceptibility to follow well learned routines, Low ET systems tend to perceive the outside world as more noisy and volatile and therefore seem more dependent on external sources of information. In this sense, error-tolerance may be an important cognitive bias when considering the interaction of humans in ever-changing environments that require rule-based decisions. Error-tolerance seems to predict different types of mistakes (e.g. perseverative errors or lost-shift errors) and recovery processes and is associated with specific patterns of EEG activity during the anticipation and evaluation of error feedback information.

We believe, the present results not only contribute to a better understanding of the mental processes underlying human fallibility in error-prone environments, but also point to the importance of taking into account individual differences related to the monitoring of complex chains of commands (Norman, 1988; Reason, 1990) when examining and judging different mistakes performed by human operators in their interaction with real-working environments.

## 7

# Discussion

## DISCUSSION

Complex behavior implies the evaluation of behavioral progress and the flexible adaption to changing contingencies in accordance with specific context goals. A key component of adaptive behavior, then, concerns the identification and correction of differences between intended and consequent executed actions, or errors.

The study of the neurophysiological bases of human error-monitoring has advanced enormously throughout the last two decades. Since the seminal work by Patrick Rabbitt in the mid-60s showing that people naturally do monitor their performance and, therefore, are likely to correct and compensate their errors, a plethora of studies have been influential in characterizing the brain networks and neurophysiological mechanisms associated with error commission, conflict detection, processing of unexpected or surprising events, and the consequent implementation of regulatory and adaptive behaviors. This field has burgeoned and provided the grounding for a very interesting research program regarding high-order cognitive control, decision-making and learning processes.

In the present thesis I tried to contribute to this promising research field by addressing novel questions and bringing new ideas to study well described neural mechanisms and temporal dynamics of error-monitoring in more extended contexts in which humans commit errors. The main intention, therefore, was to provide a broad and inclusive framework regarding the electrophysiological mechanisms involved in many aspects of action monitoring during goal-directed behavior. I believe that the questions that have been addressed here are of relevant importance to understand the extent and depth to which brain error-monitoring mechanisms can be explored from an expansive ecologic perspective.

Throughout my research I have used electrophysiological tools, fundamentally ERPs and time-frequency analysis, and novel experimental paradigms which involve extended interactions between the performer and the environment, require a symbiosis of multimodal processes (perception, action, attention, learning and decision-making) and the expression of complex intentions of the performer. In this chapter I will summarize and interpret the main findings from the studies that constitute this dissertation in order to accommodate the insights they may provide within a global perspective. Then, I will propose some renew ideas for future research.

## 7.1 Monitoring and regulation of conflicting unattended events mediated by medial prefrontal networks

Earlier theories postulate that the monitoring and adaption of conflicting and erroneous response tendencies is effortful and relies on top-down control and conscious processes that are mainly orchestrated by medial prefrontal neural networks (Jack and Shallice, 2001; Posner et al., 1994; Posner & DiGirolamo, 1998). However, there is substantial amount of detrimental information that is processed outside the scope of our covert attention, unconsciously, and must be evaluated and regulated on an effective and flexible manner so that appropriate behaviors can still succeed. In this regard one may ask whether the monitoring of unattended conflicting inputs can be also modulated by (top-down) control areas such as the medial prefrontal cortex (mPFC).

In the first study of this dissertation (Chapter 3) we tracked the time course of signals in the human brain related to the monitoring of unattended conflicting events by recording ERPs while participants performed a novel version of the Flanker task, in which spatial attention was manipulated to hinder the conscious access of task-irrelevant events located at parafoveal unattended locations. Our task design succeed in demonstrating that unattended conflicting stimuli, which could not be consciously assessed and available for accurate report (Lamme et al., 2003), still influenced subjects' behaviour, increasing their error rates and slowing down the speed of their responses when they correctly supressed error-prone conditions.

More important, the mPFC was activated by unattended conflicting events as revealed by an increase of theta-related ERP activity and theta power in medial-frontal areas right before the onset of correct, yet error-prone, responses. A similar pattern of medial-frontal theta activity was also observed in response to attended conflicting events replicating previous findings showing theta-related medial-frontal activity triggered by fully attended foveal conflicting stimuli (Nigburg et al., 2012; Cavanagh et al., 2012). Using source localization methods it was confirmed that the observed theta-related ERP activity in both conditions and during error commission was originated in the ACC, a key structure of the mPFC crucially involved in action-monitoring and regulation (Botvinick, 2001; Holroyd et al., 2004; Ridderinkhof, 2004). In this sense the processing of attended and unattended action-conflicting information may share similar neural computations related with action-monitoring, with the distinction between the two conditions most likely arising from the quality of the representations of the conflicting stimuli.

Another important finding of this study was that the power of medial-frontal theta activity predicted the extent to which participants slowed down their responses in response to both attended and unattended conflicting events, suggesting a role of theta oscillatory activity during reactive inhibitory control processes which are ensued to suppress erroneous response tendencies even in the absence of attentional (top-down) control processes. These findings are in agreement with recent research showing the influence of medial-frontal theta oscillations during the implementation of adaptive control processes in a variety of contexts involving error/conflict monitoring and regulation (Cavanagh et al., 2009; Cohen & Donner, 2013; Narayanan et al., 2014; Pastotter et al., 2013).

This study further extends previous research regarding the role of the mPFC in cognitive control by showing that this region not only detects and signal the occurrence of conflicts in information processing independently of the influence of top-down control processes, as pointed by the conflict monitoring theory (Carter et al., 1999; Botvinick et al., 1999), but also has a regulative function.

Moreover, our data suggest that although early mechanisms of covert attention influence the likelihood of conscious recognition (Cohen 2012; Lamme, 2003), unattended (unconscious) visual information can still be causally efficacious in affecting behavior and influencing high-order neural networks of cognitive control. To a certain extent these results are in agreement with recent findings suggesting an overarching role of medial-frontal networks in several strategic forms of cognitive control, such as error-monitoring (Hester, Foxe, Molholm, Shpaner, & Garavan, 2005; Nieuwenhuis et al., 2001) and the inhibition of prepotent erroneous actions (Sumner et al., 2008; van Gaal et al., 2008;) that are triggered by events that cannot be consciously assessed.

In sum, this study was important to demonstrate that a substantial amount of visual information that is processed outside the focus of overt attention and not available for conscious access activate mechanisms in medial prefrontal control networks related to high-level cognitive control processes. Moreover it showed that neural oscillatory theta activity in the mPFC may be responsible for the monitoring and regulation of potentially inappropriate actions that are automatically triggered even by environmental events to which we remain oblivious, extending current findings on the role of frontal theta oscillations in cognitive control and on-going research showing the influence of unconscious processing in high-level cognitive control processes.

## 7.2 Error-monitoring and error-awareness mechanisms during the acquisition of motor skills

A critical aspect for learning is how we use the information from our errors to improve performance. There is a great wealth of evidence regarding the brain mechanisms involved in error processing; however less we know about how these mechanisms are engaged during the acquisition of new motor skills. Concerned with this issue, in the study 2 (chapter 4) we inspected how the acquisition of a new skill lead to changes in cortical brain activity related to error-monitoring and error-awareness. To accomplish this goal, we recorded the EEG activity of musical untrained individuals while they learned a series of musical rhythmic patterns. Improvements in the skill to learn musical rhythmic patterns were accompanied by transitory ERP changes signaling distinct brain states commonly associated with automatic error/conflict monitoring and error-awareness processes.

The results from this study revealed a special engagement of medial-frontal cortex error-monitoring mechanisms, reflected by a greater activity of the ERN component, in early stages of rhythm learning. This finding demonstrates a greater engagement of medial-frontal regions tracking a large number of inappropriate response tendencies, monitoring competition between multiple conflicting motor plans and, consequently, signaling the need for increased attentional control during this learning stage (Botvinick et al., 2001; Carter et al., 1998; Nachev et al., 2006; Yeung et al., 2004). These results are in line with previous neuroimaging studies showing the prominent role of medial-frontal networks during early stages skill learning signaling the need for the implementation of top-down control processes when task performance is more challenging (Jenkins et al., 1994; Jueptner et al., 1997; Ramnani and Passingham, 2001; Toni et al., 1998).

In contrast, electrophysiological signatures predicting error-awareness developed only during later stages of learning, when the representations of the target rhythm were stronger and, consequently, errors became more salient to the performer. At the neurophysiological level, the emergence of error-awareness was identified by the increase of a centroparietal positive ERP signal resembling the P3b and the late Pe components in later stages of learning. At the neurophysiological level our results point to dissociation between error-monitoring and error-awareness processes in the human brain, which is consistent with previous research in error processing (see also Endrass et al., 2005; Overbeek et al., 2005; Steinhauser and Yeung, 2010). Remarkably, this centroparietal positive ERP was enhanced in participants who reproduced the rhythmic patterns with higher proficiency. This finding may indicate that during skill acquisition conscious error evaluation, or the metacognitive process leading to the subjective experience of error-awareness (Yeung & Summerfield, 2012), is contingent with the status of learning of the performer and consequently depends on the consolidation of internal representations of the task on-hand. This idea gives further support to other evidences relating centroparietal EEG activity with subjective response certainty (Hillyard et al., 1971; O'Connell et al., 2012). Accordingly neural activity echoes the accumulation of internal evidences underlying the awareness of actions or stimuli which are motivationally relevant and which constitute the formation of decisional processes related with judgements of confidence and accuracy of given response (Steinhauser & Yeung, 2010; Yeung & Summerfield, 2012).

This research extended previous work on error processing mechanisms in simple reaction time tasks to more ecologically valid contexts that mirrors many situations in which we acquired new skills. Moreover, it offered a possibility to merge different fields of research, the emerging field of Music Cognition and Neuroscience, the field of Motor Control and Learning and the field of Action Monitoring and Cognitive Control, which have been developed independently but, inherently, are interrelated.

## 7.3 Self-generated errors and errors reflecting agency violations are detected by distinct neural networks

Normally, the actions that we perform with our body are attributed to our own agency and therefore we feel in control over those actions. Yet, there are intriguing clinical cases of individuals experiencing abnormal experiences of agency reporting, for example, that their hand often moves on "its' own accord" as if controlled by "alien agents". These bizarre and remote cases may let us enquire, then, whether our brain is able to distinguish erroneous-type of actions that are self-generated (e.g. pouring salt in a coffee cup as result of a distraction) from errors that are triggered without ones' will (e.g. my hand suddenly grabs the salt without my intention and pours it in the coffee cup).

In the study 3 (chapter 5) of this Thesis we invented a pioneering experimental situation in which we recorded the EEG activity of healthy participants embodied into avatars in order to disentangle this question. Using ERPs we compared neurophysiological signatures related to self-generated motor errors and errors which were imposed by an "alien-agent" without subjects will.

Our findings showed that self-generated and alien-errors were associated with very distinct ERP signatures - different spatial cortical activity and different processing latencies. These results highlighted the existence of two independent, but not exclusive, error-monitoring systems in the human brain which are responsible for the monitoring of the correctness of our actions in accordance with our goals and the evaluation of our sense of control over those actions, or our sense of agency.

We observed that self-generated errors triggered a classical frontocentral ERN. As discussed in chapter 1 of this dissertation it is likely that the ERN response reflects the operation of a rapid internal-error monitoring system in the medial-frontal cortex, which uses internal forward models of the motor plan (efference copy) to detect errors in on-going actions (Coles et al., 2001; Gehring, Gross, Coles, Meyer, & Donchin, 1993; Holroyd, Yeung, Coles, & Cohen, 2005; Rodriguez-Fornells, Kurzbuch, & Munte, 2002). Accordingly this system aids the implementation of fast corrective and compensatory behaviors in an automatic fashion without the need of relying on sensorial feedback information.

Interestingly, the insertion of alien-errors on participants' correct movements, which impaired their natural sense of agency, elicited a delayed negative ERP signal (at about 400 ms) with a parietal distribution. This ERP response could reflect the output of an external-error-monitoring system when registering the occurrence of agency violations or the loss the sense of motor control, which at the neurophysiological level is unrelated to medial-frontal brain networks involved in action-monitoring processes such as error commission, error observation, or error feedback evaluation (see for a review paper Ullsperger et al., 2014). Accordingly, this external monitoring loop may use internal forward model copies of the current motor command to predict the feedback that the action will produce. If the comparison between predicted and actual feedback generates no error, then no feeling of losing action control is perceived. As this external error-monitoring system, relies ultimately on the processing of different re-afferent feedback information, arriving at the somatosensory, visual and auditory regions with different neural delays, we observed that the timing required by the human brain to detect errors or violations between one's intentions and it's actual consequences (i.e. based on the feedback information) may have a delay of 350-400 ms. This suggests that the detection of a loss of agency is slower than the detection of self-generated errors which depends exclusively on the central efference copy. Moreover this operation was not computed in the medial-frontal cortex, but on parietal regions. The parietal distribution of the negative ERP signal following agency violations converges with previous functional neuroimaging and lesion studies in which the importance of the posterior parietal regions has been highlighted in relation to diminished feeling of agency (Farrer et al., 2003; Farrer, Bouchereau, Jeannerod, & Franck, 2008) and neural computation processes between motor intentions and action consequences (Desmurget et al., 2009; Sirigu et al., 2004).

Remarkably, the amplitude of the neurophysiological signal of agency violations was correlated with the subjective feeling of body ownership with the avatar, which suggests a relation between the internal representations of the body image towards the self-represented avatar and the impact that unexpected imposed actions by the self-represented agent has on one's natural phenomenal experience of agency. This result is important as it points out the relevance of considering one's natural body image during the evaluation of one's own 'action semantics'. This result may also be relevant to previous studies on agency attribution which have used tasks wherein subjects judge whether they did or did not cause a specific external event (e.g. while playing a video game they judge whether they caused the joystick movement shown on the video or not) (e.g. Farrer et al., 2008), which *per se* say little about the phenomenology or experience of agency within one's own body.

Intriguingly, the negative ERP following alien-errors showed a strong similarity in topography and latency to other electrophysiological signatures related to semantic or conceptual violations, specifically the N400 component (Kutas &

Federmeier, 2011). Thus an interesting question that arises here is the extent to which observing the representation of an embodied body performing a non-planned, incongruent or unexpected, action might be evaluated in a similar fashion as a semantic-conceptual violation.

As this neural signal of agency violations may result from an internal conceptual clash between the predicted consequences of ones' current action (I move my hand to thee left) and the actual feedback resulting from that action (I saw my hand moving to the right), it is conceivable to think that at the neural level this EEG signal could share similar characteristics with other error-prediction EEG signals in the human brain which have been widely described in the literature as indexing discrepancies between actual and intended (predicted) responses or feedbacks (Holroyd & Coles, 2002). According to previous theories of error processing, the insertion of alien-errors should result in a retrospective medial-frontal ERP signal (e.g. the feedback error-related negativity). However, in the present study agency violations did not trigger such type of retrospective medial-frontal ERP signal, leaving open for future research the relation between medial-frontal networks involved in action monitoring and more posterior parietal brain networks underlying the sense of control over voluntary actions.

The study 3 of this Thesis provides crucial new evidence about how to distinguish at the neurophysiological level self-generated errors from errors imposed by "alien agents" and revealed, for the first time, an electrophysiological signature specifically related to agency violations. These findings shed new light on the timing required by the human brain to differentiate self-generated and alien-errors, providing new neural evidences regarding the integration of internal and sensory feedback information in the build-up of a coherent sense of agency during action monitoring. Furthermore the new experimental setting presented here offers a unique opportunity to study the experience of agency while being embodied actors of our own actions, opening new avenues on current research focused on the mental processes underlying agency attribution in healthy and clinical populations.

## 7.4 Electrophysiological evidences of individual differences in Error Tolerance in humans

Humans differ considerably in the way they interpret and appraise their errors and error feedback information about their performance. These differences are fundamental to understand intra-individual variation during the initiation and maintenance of cognitive control and to rationalize about decisional processes carried out by humans in complex and uncertain environments in which errors are likely to take place.

Therefore, in the last study of this Thesis (study 4, chapter 6) we shifted the focus of attention from the diversity and range of error-monitoring processes to explore the intrinsic variability of those processes during goal-directed decision making routines. In the study 4 we investigated whether individual differences in error-tolerance, i.e. the tendency (or bias) of humans to follow well-programmed action plans (e.g. well-known routines) in the presence of conflicting or redundant error feedback information without the need to drastically change pre-selected courses of action, could be reflected in specific neurophysiological mechanisms associated with discrete cognitive and motivational states.

To assess neurophysiological differences in error-tolerance we examined the ERP responses of two extreme error-tolerant groups (categorized as Low and High error-tolerant) which during the performance of a learning task involving rule-based decisions showed drastic differences regarding the propensity to maintain new learned rules after receiving misleading or redundant error feedbacks. Differences in error-tolerance were characterized by different type of rule-based mistakes and specific patterns of EEG activity during the anticipation and evaluation of error feedback information. Participants with higher proneness to maintain new-learned rules, i.e. with higher error-tolerance, showed reduced attentional and motivational engagement during the anticipation of incoming feedbacks informing about their performance. This was evidenced by reduced amplitude of the SPN component, an electrophysiological signal sensitive to anticipatory attentional and motivational states to informative feedbacks (Brunia et al., 2011; Fuentemilla et al., 2013; Moris et al., 2013). Participants with higher error-tolerance also showed decreased cortical responses during feedback processing, as evaluated through the FRN and P3 components (Fischer & Ullsperger, 2013; Yeung et al., 2005). Thus, individuals whose decisions rely mostly on internal-(rule)-based knowledge seem less responsive to and less dependent on external feedback information, operating predominately in a feedback-independent mode. In contrast individuals with Low error-tolerance were more prone to change frequently their behaviors after receiving error feedbacks and directed an increased attention to external feedback information.

Nevertheless, is spite of individuals with High and Low error-tolerance were differently affected by misleading or redundant error feedback information, when they collected enough evidences that new task contingencies implied the reprograming of pre-selected courses of action, their neural patterns of EEG activity predicting rule-based behavioral adjustments converged to a common path, suggesting common adaptive cortical mechanisms of updating of feedback information regardless to the error-tolerance bias.

The results from this study support the existence of specific cognitive and motivational processes underlying individual differences in error-tolerance among humans, which may bias the commission of different type of deliberate erroneous actions during complex decision-processes.

These differences seem to be related with different modes of cognitive control in performing tasks involving the monitoring of conflicting, redundant, and errorful information; as a matter of fact, individuals with high-error-tolerance seem to act, predominantly in a feedback independent mode, while subjects with low error-tolerance tend to give more importance to external feedback information when guiding their decisions in more uncertain environments.

In this sense, if common sense suggests that humans have different tolerance or reactivity to negative (errorful) feedback information, we believe that these findings are relevant and stress the importance of taking into account individual differences related to the monitoring of complex chains of commands (Norman, 1988), contributing to the on-going research focused on understanding the mental processes behind human fallibility in error-prone scenarios.

## 7.5 Interesting questions for future research

This dissertation addresses a number of novel ideas regarding the neurophysiological mechanisms underlying action-monitoring and adaptive behavior in a broad set of contexts. The novelty of the findings described here, though, may also open up a wide range of interesting questions for future research which may help to prove or disprove some of the ideas that have been forward in the last section as well as enrich our understandings regarding the nature of the brain mechanisms and cognitive processes underlying adaptive and goal-directed behavior. In the study 1 it was proposed that a great amount of relevant unattended information, which remains at some level unconscious, still modulates the activity of prefrontal neural networks in charge of detecting and overcoming potential disadvantageous and inappropriate actions. An interesting question that still remains to be addressed, though, is the extent to which unattended inputs which are relevant to on-going performance (take the example of the dog that is about to jump on the street) can thereby influence early stages of visual processing and consequently modulate prefrontal neural networks in a dynamic manner changing the focus of covert attention towards the relevant stimuli. In other words, can conflict monitoring operate already at early stages of visual processing (in the visual cortex), transmitting an error signal to prefrontal networks enhancing on-flight top-down visual attention for prioritizing unattended inputs according to its 'relevance'? If so, are these mechanisms also orchestrated by theta oscillations?

Patients with spatial neglect have been described as being unable to respond, consciously orient to and report stimuli located in the contralesional visual space, even though they are not blind to stimuli on that side (Husain & Rorden, 2003). However, are they able to monitor the presence of relevant unattended visual inputs from the neglected visual space and therefore adapt their behavior accordingly in a goal-directed fashion? If so, could we observe in those patients behavioral interference effects caused by neglected conflicting stimuli and related dynamic engagement of medial prefrontal areas associated with action monitoring and regulation?

Study 2 offered a broad picture of how brain mechanisms related to error-monitoring and error-awareness processes are engaged during the learning of a new skill. Given the hypothetical role of the striatal dopaminergic circuitry in error-monitoring and cognitive control processes (Holroyd & Coles, 2002) as well as in motor procedural learning and skill acquisition (Barnes, Kubota, Hu, Jin, & Graybiel, 2005; Kawashima et al., 2012; Yin et al., 2009), to what extent changes in striatal dopamine release can influence positively (or negatively) error-monitoring and error-awareness processes during skill learning? Besides, what is the role of oscillatory theta activity synchronizing action-monitoring (in the MFC) networks with control (lateral PFC) and motor networks throughout the process of skill acquisition? Does the strength of theta-band phase-synchronization between midfrontal and control and motor networks influence positively skill learning processes? Regarding the specific task of rhythm learning should we expect differences regarding the electrophysiological activity related to error-monitoring and error-awareness between professional and untrained musicians? These are interesting new questions raised by this experiment.

Study 3 put forward for consideration a dissociation between neural networks involved in action monitoring, in the medial-frontal cortex, and networks underlying the sense of control over our own actions, in the parietal cortex. A pertinent question to be tested in the future is if these brain networks interact dynamically and in a directionally specific manner, and whether they use different frequency bands to support our natural phenomenal experience of agency during voluntary movement execution. Furthermore, in the study 3 we combined last technological improvements in virtual environment experimentation with time-sensitive neuroimaging measures to disentangle the brain mechanisms underlying error-monitoring and agency. This new experimental approach offers a unique opportunity to study the experience of agency while being embodied actors of our own actions; thus it may motivate future research concerned with impairments of agency attribution in clinical populations, for instance, schizophrenic individuals reporting delusions of control, auditory hallucinatory experiences or thought insertion (Frith et al., 2000; Frith, 2005). In this sense could we use the parietal ERP signal that we observed following agency violations as a neural marker for delusion experiences in such patients?

Assuming that individuals with high and low error-tolerance seem to be differently affected by external feedback information, favoring feedback-independent vs. feedback-dependent modes of control, to what degree can these differences be reflected in micro-structural brain differences and functional hemodynamic responses related to reward processing brain networks (Camara, Rodriguez-Fornells, & Munte, 2010)? Are persons with low error-tolerance more prone to perceive external feedbacks as inherently more aversive to some degree? To which degree the error-tolerance trait can regulate which cognitive modes of control, reactive vs. proactive (Braver, 2012), are preferred during the performance of complex tasks with high cognitive control demands? Finally, framing the concept of error-tolerance into principles of reinforcement learning (Behrens et al., 2007; Jocham et al., 2009), on interesting question is to which extent the error-tolerance bias reflect deficits in reward prediction error mechanisms and/or dynamic processes of updating of contextual information over time (i.e. learning rate).

## 7.6 Other reflexive questions regarding the present studies

Approaching the study of human performance in more ecologic and realistic environments is of crucial importance to better understand the neural and cognitive bases underlying imperfections in action. This approach has important implications for future research, but, more importantly, contributes to prevent and mitigate negative consequences that human failures cause at the individual and social level. Moreover, such approach may also open new possibilities to improve the quality of life of many individuals who are prone to commit errors in many tasks of their everyday life or who have severely impaired their capacity to monitor their actions in a goal-directed fashion. In continuation some reflections would be shared with the reader regarding the implications of approaching the study of human errors from more ecologic perspectives.

Taking into account that there are people with different tolerance to external errorful feedback is of vital importance for different professionals who are interested in improving human performance in a wide set of working contexts. For instance, the conceptualization of an error-tolerance trait-like factor, which echoes specific cognitive and motivational processes, may be important in the development of new psychological tests and tools to evaluate the abilities of operators applying for job positions requiring the monitoring of error feedback information in complex technological systems. In this case, psychologist may be better prepared to select operators who are less likely to commit errors in stressful or uncertain environments, or even improve their cognitive abilities to cope with error-prone environments or motivate them to perform better. These considerations have also important implications regarding the social impact that individual errors can have.

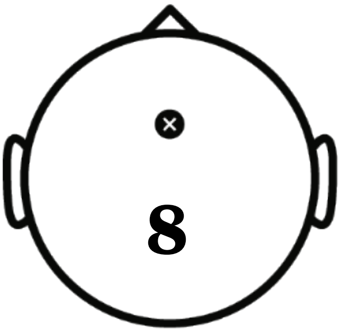
Furthermore, as a proper designed system must take into account the properties of the people who use it, the consideration of factors such as individual differences in error-tolerance, might be of importance for the design of complex systems operated by humans (e.g. aircrafts, nuclear power plants). Therefore, we can redesign the human-machine interface so that its behavior can be shaped to the human characteristics in appropriate ways. In this regard, there may be also the case of implementing artificial intelligence interfaces improving automation in order to reduce errors caused by human factors. In fact, the improvement of mechanical engineering along with our current knowledge regarding the biological bases of human action-monitoring is an important step to improve human performance and to alleviate the consequence of errors.

The identification of neural signals of action-monitoring in more realistic scenarios may allow the development of brain-computer interfaces (BCI) that use error-related brain signals to improve individuals performance in a wide variety of cognitive demanding contexts. For example, EEG-based BCI interfaces may be used in the future to improve the prediction of driver's intended actions in intelligent cars to tune their driving assistant systems to improve their performance. Moreover, the development of such interfaces may also bring fundamental therapeutically benefits for populations affected by pathological or declining performance monitoring (e.g. Parkinson patients) or individuals with impairments in the ability to perform volitional movements (e.g. people with paralysis). The development of biomedical and biotechnology research is already a reality in restoring the mobility and independence for people with paralysis or without limbs by translating action-related neuronal activity directly into control signals for assistive artificial devices (e.g. robots or prosthesis).

However, the possibility of using artificial intelligence interfaces or to control other bodies or body parts (e.g. avatars and robot devices) may also confront us with important issues concerning the moral and legal status of our actions. In this last case, to what extent we may be legally responsible for actions that our surrogate body (e.g. an avatar controlled by ourselves) take without our intention (for example due to a technical malfunction)? And to what extend our feeling of agency and legal consequences would be the same when we have different bodies, acting in different places?

Furthermore, some of the findings reported here also suggest that many of our everyday goal-directed behavior unfolds entirely automatically without requiring much conscious or voluntary attentional control. Though, this assumption also brings to mind the question whether all cognitive and neural operations can be initiated unconsciously, whether this might only be the case for highly trained and over-learned behaviors and whether conscious and unconscious processes may share similar neural mechanisms but varying in the quality of the representations.

Finally I hope that the research program presented here proves the good applicability of cheap neuroimaging techniques such as the EEG which can provide reliable good-time measures to explore the functioning of neural networks underlying human performance monitoring in more ecologic scenarios.



# 8

# Conclusion

## CONCLUSION

In the present thesis I developed new conceptual toolkits and offered a new set of multilevel constraints for current concepts of error-monitoring in cognitive neuroscience. The conclusions drawn from this work may shed some light on various aspects of the brain mechanisms underlying error-monitoring and cognitive control processes, but also, open new avenues for future research in extended ecological contexts.

This work brings together different lines of research that in some of the cases had progressed independently of the current research approaches devoted to the study the neurophysiology of human error processing. Therefore, considering the possibility of studying errors in extended and complex contexts, the present work embraced hotly debated issues in the field of Cognitive Neuroscience, such as: (i) the role of attentional and (un)conscious processes in higher cognitive control functions; (ii) performance monitoring and skill acquisition; (iii) agency, the self and action; and (iv) individual differences during decision making processes.

Based on the experiments reported in this dissertation the following conclusions can be drawn:

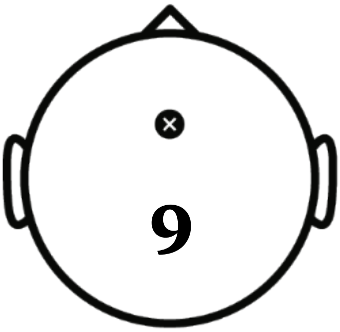
1. There is a substantial amount of visual unattended information to which we remain oblivious but yet influence our behavior. Stimuli that we do not attend and remain unconscious, still modulate prefrontal control networks (as the medial prefrontal cortex), traditionally assumed to be exclusively associated with conscious and top-down control processes. Neural theta oscillatory activity may stand as a plausible neurobiological mechanism by which medial-frontal networks monitor and regulate inappropriate actions that are automatically triggered by conflicting environmental information.
2. Practice leads to functional changes in the brain mechanisms associated with error-monitoring and error-awareness during the acquisition of new motor skills. At early stages of skill learning a greater engagement of error-monitoring mechanisms of the medial-frontal cortex are fundamental to overcome a higher number of erroneous and conflicting motor plans. Error-awareness, in turns, seem to depend on the strength of the representations associated with the task on-hand which are refined with practice, such that, error-awareness is intimately associated with the status of learning of the performer.
3. Self-generated errors and errors related to agency violations are associated with distinct error-monitoring systems, which together are involved in providing a coherent sense of the agency of our actions. Self-generated errors rely exclusively on an internal error-monitoring system operating in the media-frontal cortex. Errors that reflect agency violations, however, require the participation of an external error-monitoring system, which is

slower and operates in parietal regions of the brain. This system is responsible for checking whether external feedback information (multimodal re-afferent feedback information) following a given motor action matches the predicted body state so that the cognitive agents may feel in control over their actions. In this sense, prior semantic representations of the body state (sensory predictions) have to be coherent with resulting state of the body, for us to feel that we are causally responsible for a given action.

4. Different thresholds of error-tolerance in humans are related to different type of rule-based mistakes following error feedback information and distinct patterns of cortical activity during the anticipation and evaluation of redundant error feedback information, which expose discrete cognitive and motivational brain states to cope with errorful information. These findings stress the importance of taking into account individual differences related to error processing and monitoring of complex chains of commands, to understand the mental processes behind human fallibility in uncertain and ever-changing environments.

I believe that the findings from these investigations and the ideas present along this dissertation may offer a great contribution to the ongoing research in cognitive neuroscience regarding action monitoring and adaptive control of behavior.

Moreover, these findings may be of great importance to validate current neurophysiological evidences and theories regarding human error processing and cognitive control processes, but fundamentally, may offer new insights to understand the extent and depth to which the human error-monitoring system can be studied in extended and ecological contexts that mimic the complex cognitive demands of everyday life in which human agents have been find to commit errors.



# 9

# Spanish Summary

## 9.1 Introdución

### 9.1.1 El estudio de los errores humanos en la Neurociencia Cognitiva

El error forma parte de nuestras vidas. No obstante los seres humanos aprendemos de ellos y guiamos nuestras acciones monitorizando dichas imperfecciones y ajustando nuestra conducta de una forma flexible y adaptativa. De esta forma, la adaptación al ambiente requiere de sistemas cerebrales que permitan monitorizar las consecuencias de nuestras acciones y, en caso de ser necesario, modificar las acciones que han comportado resultados no deseados.

El estudio de los mecanismos cerebrales asociados a la monitorización de nuestras acciones, y nuestros errores, se ha desarrollado de forma muy importante durante los ultimos 20 años y consituye un importante campo de investigación de la Neuroriencia Cognitiva. Nuestro entedimiento sobre estos mecanismsos se ha desarrollado principalmente debido al descubrimiento de respuestas electrofisiológicas (estudios de electroencefalografía, EEG) y redes neuronales que son sensibles a la comisión de errores, al procesamiento de eventos negativos o indeseados y a la evaluación de información conflictiva. Estos hallazgos han desvelado varias dinámicas relacionadas con los procesos de control cognitivo, toma de decisiones y aprendizaje orquestados por nuestro cerebro; y por lo tanto han mejorado nuestro entendimiento sobre como el cerebro controla nuestras conductas y permite adaptarnos al entorno en función de nuestros objectivos.

No obstante, aún queda un largo camino que percorrer. Un objetivo crucial de esta rama de la neurociencia es el desarrollo de métodos que permitan estudiar los procesos mentales inherentes a la monitorización y regulación de nuestra conducta en contextos con mayor validez ecológica. Fundamentalmente, el primer paso consiste en poner a prueba hipótesis y modelos que ya conocemos sobre el procesamiento de errores por el cerebro en experimentos más realistas que imitan con exactitud la complejidad de varias exigencias de la vida cotidiana.

### 9.1.2 Paradigmas experimentales tradicionales para el estudio de errores

Los paradigmas clásicamente usados para estudiar los procesos cognitivos y cerebrales inherentes a la monitorización de errores se asimilan por el uso de condiciones experimentales que generen conflicto o interferencia cognitiva durante la generación de una acción. A menudo, este conflicto es inducido por eventos perceptivos que desencadenan tendencias de respuestas inadecuadas que necesitan ser corregidas o suprimidas de forma flexible para evitar errores. Estas tareas usualmente requieren la realización de acciones sencillas; como pulsar un botón en respuesta a eventos visuales o auditivos (por ejemplo, responder de forma muy rápida a la dirección de flechas presentadas en una pantalla de ordenador: “contesta izquierda o derecha"). Ejemplos de estos paradigmas son la tarea "Erikson Flanker" (Eriksen y Eriksen, 1974) o la tarea de "Stroop" (Stroop, 1935). La lógica de estas tareas es que hay condiciones de elevada incongruencia (entre percepción-acción) que generan automáticamente conflictos en el plan de acción, aumentando la tendencia a ejecutar respuestas erróneas. A nivel conductual, las condiciones con mayor conflicto están asociadas a mayores tasas de error y a tiempos de reacción más lentos (Botvinick et al., 2001). En estos contextos, los errores usualmente reflejan acciones no deliberadas (*'action slips'*) (p. ej.: echar sal en una taza de café como consecuencia de una distracción), como resultado de lapsos atencionales, perceptivos o pierdas momentáneas de control cognitivo (Botvinick et al., 2001; Eichele et al., 2008; O'Connell, 2009).

Sin embargo, este tipo de tareas no reflejan la complejidad de las demandas cognitivas que los seres humanos requieren en su vida cuotidiana. En realidad, fuera del laboratorio no todos los errores son el resultado de fallos de atención o lapsos, que conllevan a fallos en tareas tan sencillas como las mencionadas anteriormente, pero pueden estar asociados a una larga escala de factores. Por lo tanto, es importante desarrollar nuevos paradigmas experimentales con el fin de investigar los errores en contextos de mayor validez ecológica.

### 9.1.3 Indicadores cerebrales asociados al procesamiento de errores en humanos y su corrección

En la década de los 60, Patick Rabbit especuló por primera vez sobre la importancia de un sistema responsable por detectar y regular errores en nuestra conducta. Basado en estudios conductuales (Rabbitt, 1966a; 1966b; véase también 2002) el demostró que tendemos a corregir nuestros errores de forma casi automática (y quizás inconsciente en varias situaciones) y que normalmente ajustamos nuestra conducta de forma efectiva y adaptativa después de la comisión de errores, por ejemplo, enlenteciendo la velocidad de nuestras respuestas o aumentando la precisión de ellas.

De alguna forma, estos resultados dejan antever la existencia de un sistema cerebral de monitorización de acción que ayuda a la implementación de conductas compensatorias y adaptativas en respuesta a errores, situaciones de gran conflicto, novedad y dificultad de la tarea. Todas estas situaciones, de alguna forma, requieren un cierto grado de control cognitivo para que el organismo detecte discrepancias entre los objetivos de sus acciones y sus respuestas, teniendo en dichos casos que dirigir la atención de cara a poder inhibir la acción en curso (*si aún fuera posible*), corregirla o bien intentar evitar en el futuro cometer dicho error de nuevo.

Algunos de los resultados más relevantes en el estudio de dicha monitorización se han llevado a cabo mediante el estudio de la actividad eléctrica cerebral (electroencefalografía), concretamente mediante la técnica de potenciales evocados

(PES). En concreto se han descrito respuestas cerebrales que dan cuenta de mecanismos cerebrales de control sobre las acciones realizadas, señalando errores, eventos conflictos, inesperados o negativos que necesitan ser compensados y regulados de una forma flexible para que nuestra conducta sea la más apropiada en función de nuestros objetivos.

En primer lugar, existiría un mecanismo de control interno que regularía la discrepancia entre una respuesta correcta y una respuesta incorrecta. Dicho mecanismo de control se vería reflejado en una respuesta eléctrica cerebral (la negatividad asociada al error, *error related negativity, ERN,* (Falkenstein et al., 1990; Gehring et al., 1993) que aparece 0-100 ms después de cometer un error y que presenta una topografía frontocentral negativa. Por otro lado, existiría un mecanismo de control externo que se llevaría a cabo mediante la evaluación de los resultados de las acciones. Dicho mecanismo se vería reflejado en los potenciales evocados por el componente Negatividad asociada al Feedback (*Feedback-Related Negativity, FRN,* Holroyd & Coles, 2002; Miltner et al., 1997*)* que aparecería 200-400 ms después de la presentación de un estímulo informando de una consecuencia negativa o inesperada de la acción realizada. La observación de errores de otros agentes también suele desencadenar un PE fontocentral similar a la FRN (Van Schie et al., 2004). Hay aún otro componente negativo frontocentral que está asociado a estímulos conflictivos que requieren la supresión de respuestas potencialmente erróneas y que aparece entre 200-300 ms después de la presentación de dicho estimulo conflictivo y justo antes de la ejecución de respuestas correctas pero muy propensas al error: el componente *N2* (van Veen & Carter, Yeung et al., 2004). El análisis de las fuentes generadoras de la ERN, FRN y N2 muestran un origen común de estas señales neuronales en áreas de la corteza medial prefrontal (mPFC), sobre todo el anterior cingulado (ACC) y área (pre) suplementaria motora (preSMA y SMA) (Yeung et al., 2004; Holroyd & Coles, 2002; Grundler et al., 2013).

Más recientemente estudios interesados con el carácter oscilatorio de la actividad eléctrica cerebral - que permiten estudiar la actividad eléctrica cerebral en diferentes frecuencias sin perder la alta resolución temporal que permite la electroencefalografía - han demostrado que errores y eventos negativos o conflictivos están asociados con respuestas cerebrales de baja frecuencia en el rango de theta (4-8 Hz) (Cavanagh et al., 2009: Cavanagh et al., 2012; Marco-Pallares et al., 2008; Nigburg et al., 2012).

Por otra parte estudios realizados con resonancia magnética funcional (*functional magnetic ressonance imaging*, fMRI) en paradigmas dónde se generan errores (por ejemplo, usando tareas de Eriksen Flanker o Stroop) han demostrado la implicación del mPFC durante la generación de errores o acciones que son potencialmente relacionadas con acciones conflictivas y erróneas y eventos negativos e inesperados (Carter et al., 1988; Jocham et al., 2009; Marco-Pallares et al., 2008; Ullsperger & von

Cramon, 2001). Estos estudios también han demostrado la implicación de otras áreas como la corteza dorsolateral prefrontal, la corteza inferior frontal y la corteza insular.

Conjuntamente, estas evidencias, sugieren que el mPFC es un importante locus de evaluación de nuestras acciones, siendo especialmente sensible a la computación de errores. Por otra parte, esta región no sólo participa en los procesos de evaluación de la acción, pero también hay varias evidencias que indican la participación de esta área en varios aspectos reguladores de nuestras acciones (por ejemplo, corregir errores y ajustar la conducta después de errores o eventos conflictivos y demandantes) (Kerns et al., 2004; Garavan, 2002; Marco-Pallarés et al., 2008; Danielmeier et al, 2011). De hecho, lesiones en esta área comprometen de forma grave la capacidad de uno adaptarse de forma efectiva y flexible a desafíos del ambiente (Ridderinkhoff et al., 2004; Shett et al., 2012).

De acuerdo con teorías influyentes de control cognitivo, el mPFC, en respuesta a eventos conflictivos y errores, alerta al sistema para aumentar el control y guiar los ajustes de comportamiento - aumentando la atención hacía la tarea o informando las áreas motoras del cerebro para corregir o ajustar cursos de acción inapropiados (Botvinick et al., 2001; Holroyd & Coles, 2002; Kerns et al., 2004; Ridderinkhoff et al., 2004). También se ha demostrado que las señales de esta área pueden ejercer control sobre regiones del cerebro responsables del procesamiento sensorial y motor con tal de mejorar la ejecución de una tarea(Danielmeier et al, 2011;. King et al, 2010).

Varios modelos computacionales y neurocognitivos se han postulado recientemente para entender cómo el cerebro monitoriza y activa los procesos de control cognitivo necesarios en situaciones novedosas o ante la posible comisión de errores. En dicho contexto, estas teorías han delimitado el mPFC en dichos procesos ejecutivos, aunque aún se desconoce el rol funcional exacto de esta área.

La primera teoría explicativa del procesamiento de errores (‘*The Comparator Model*’) propuso que el procesamiento de errores implica un proceso de comparación entre la representación de la intención de una acción y el resultado actual de esa respuesta. En este marco, la señal ERN en respuesta a errores, y relacionada con la actividad del mPFC, refleja la respuesta de un sistema de monitorización que detecta una falta de coincidencia entre la respuesta prevista (correcta) y la respuesta errónea real, según lo determinado por el estado actual del sistema después de que la respuesta se ejecuta (Coles, Scheffers, y Holroyd, 2001; Falkeinstein et al, 1990; Falkeinstein, Hoormann, Cristo, y Hohnsbein, 2000; Gehring et al, 1993; Scheffers y Coles, 2000). Debido a que la ERN es una señal extremamente rápida (50-100 ms después de la ejecución de un error) fue propuesto que el proceso de comparación no podía depender de la información sensorial o propioceptivo, ya que esta información es demasiado lenta y no estaría disponible hasta que se haya completado la respuesta; pero en su lugar se guiaría por mecanismos internos de predicción (*'eference copy'*). El modelo sugiere que durante la ejecución del movimiento si se activa una señal de error interno, el sistema puede implementar comandos inhibidores rápidos y procesos de corrección de errores con el fin de evitar un error de ser ejecutado. La dependencia de estos mecanismos independientemente de *feedback* exterior justifica de alguna forma la automaticidad y rapidez con la cual somos capaces de corregir varios errores motores sin darnos cuenta.

Un segundo modelo bastante influyente en el campo de control cognitivo es la teoría de *Monitorización de Conflicto* ('conflict-monitoring theory'). La teoría se basa en la hipótesis de que durante la ejecución de tareas no rutinarias y demandantes el cerebro aumenta su control por un mecanismo que detecta conflictos en el procesamiento de información. Basado en una serie de estudios de neuroimagen, la teoría de *conflict-monitoring* ofrece una idea integradora que reúne una serie de diferentes resultados que muestran que el mPFC participa activamente en contextos que requieren selección de acciones en situaciones de errores o de interferencia cognitiva (Botvinick et al, 2001; Botvinick et al, 2004;. Carter y van Veen, 2007). El primer punto clave de este modelo es que el mPFC es responsable de la detección de tendencias inadecuadas y conflictivas durante la generación de una acción. De esta forma, autores de esta teoría proponen que los resultados de fMRI y de PE que muestran la actividad del mPFC durante errores y situaciones conflictivas refleja el grado de conflicto motor detectado por el sistema (es decir, en un error, la competición entre la respuesta incorrecta con la respuesta correcta, que empieza a prepararse para corregir el error). La teoría sugiere, además, que después de la detección de conflicto, el mPFC envía señales a otras áreas de la red de control atencional para implementar ajustes estratégicos en el control cognitivo (MacDonnald et al, 2000; Kerns et al., 2004) que sirvan para prevenir y remediar estos conflictos.

Otro modelo bastante influyente en los últimos años es la teoría de aprendizaje por refuerzo (*reinforcement learning theory, RL*) (Holroyd & Coles, 2002). La teoría se basa en investigaciones previas que implican los ganglios basales y el sistema dopaminérgico del cerebro en el cálculo de señales de predicción de recompensa (o castigo) en situaciones de condicionamiento clásico (Schultz et al., 1997; Schultz, 2007). Según estos estudios, los ganglios basales monitorizan eventos y computan predicciones sobre si los resultados asociados a esos eventos son favorables o desfavorables. Esta señal de error de predicción se envía al sistema, informando si el resultado es mejor (error de predicción positivo) o peor (error de predicción negativo) de lo esperado. En consecuencia, estas señales de error negativo y positivo son utilizadas por los ganglios basales para actualizar sus predicciones y luego se envían al mPFC. El mPFC, un centro de monitoreo de la acción y de selección, procesaría esta información para seleccionar las conductas apropiadas para guiar el aprendizaje y comportamientos razionales (Holroyd y Coles, 2002;. Holroyd et al, 2004; Rushworth, 2004). En este marco tanto la ERN/FRN como la activación del mPFC reflejan señales neuronales que indica si los resultados de una determinada acción son mejores o peores de lo esperado, siendo su actividad más evidente en eventos inesperados. A su vez el mPFC usaría esta información (errores de predicción) para optimizar la realización de la tarea.

### 9.1.4 La necesidad de estudiar los errores humanos en contextos más ecológicos

Como se ha comentado previamente, la mayor parte de la investigación en neurociencia cognitiva sobre errores humanos se ha centrado principalmente en la observación de errores aislados de ejecución, o lapsos ('slips'), en tareas de tiempo-reacción bastante simples. Estos errores reflejan normalmente la ejecución de acciones automáticas, y sin deliberación consciente, (por ejemplo, tomar un camino equivocado durante la conducción) a causa de, en la mayoría de las veces, lapsos de percepción, fluctuaciones de atención o males ajustes en el control cognitivo.

Sin embargo, en la vida real los errores pueden asumir distintas formas y ser de naturaleza distinta, dependiendo así de los contextos en los que se encuentre el individuo. En este sentido no todos los errores son el resultado de problemas de activación de mecanismos de atención o de control del sistema nervioso. De hecho, hay muchos errores que ocurren en circunstancias en las cuales el nivel de control atencional dirigido a la tarea es el adecuado. Para ilustrar, cuando uno empieza el aprendizaje de una nueva habilidad (p. ej.: aprender a tocar un nuevo instrumento) los errores ocurren aunque la atención sea elevada y, de hecho, contribuyen al proceso de aprendizaje; independientemente del nivel de atención y control en una tarea determinada, hay errores que son debidos a fuentes exógenas en el ambiente (p. ej.: mal funcionamiento de una máquina, o accidentes causados por terceros). Además, hay varias condiciones clínicas de individuos con anormalidades en el control de sus acciones que incurren frecuentemente en errores no intencionales (por ejemplo, enfermedad de Parkinson, patologías del lobo frontal como 'utilization bevavior', el síndrome '*alien-hand*', delirios de control en condiciones psicóticas, y muchos otros).

Por otro lado, hay aún otro tipo de errores que son más relacionados con fallos en la planificación o en la formación de una intención y por lo tanto pueden ser consecuencia de procesos de más alto nivel que implican la toma de decisiones y el aprendizaje. Por ejemplo, si uno erra en la elección de un objetivo, la acción elegida para llegar a esa meta puede ser correcta o aceptable, pero ya que el objetivo planteado es incorrecto o desajustado el resultado puede no ser lo deseado.

Además, los seres humanos exhiben diferente tolerancia al error y por lo tanto pueden ser sesgados a planificar acciones erróneas de forma a afrontar contextos de gran incertidumbre que requieren la monitorización de información redundante o conflictiva del ambiente (p. ej.: un controlador de radares de navíos de guerra o controladores aéreos).

Por lo tanto, una limitación importante de la actual línea de investigación en neurociencia cognitiva es que los paradigmas utilizados para estudiar los mecanismos cerebrales subyacentes al procesamiento de errores no cubren la totalidad de la complejidad de los entornos reales o situaciones novedosas en las cuales los seres humanos están propensos a cometer errores. La investigación actual carece de la implementación de paradigmas experimentales novedosos y creativos que impliquen tareas más complejas y multimodales, interacciones más extensas entre el agente y el ambiente y la expresión de intenciones complejas del agente.

La combinación de dichos paradigmas con las actuales técnicas de neuroimagen disponibles es indispensable para caracterizar de una forma más ecológica cómo la supervisión y la adaptación del comportamiento se implementa en el cerebro en contextos más realistas que imitan con exactitud la complejidad cognitiva de las exigencias de la vida cotidiana.

Recientemente se han realizado algunos esfuerzos en este sentido y algunos estudios han investigado procesos asociados al procesamiento de errores en tareas más complejas y de mayor validez ecológica, como por ejemplo mecanografía (Logan y Crump, 2010), tocar el piano (Ruiz et al, 2009), disparar un arma (Bediou et al., 2012). Sin embargo, estos estudios representan un tímido intento en este campo de investigación, pero deben ser tomados como una inspiración para la investigación futura. La presente tesis tiene como principal motivación seguir los pasos dados por estos estudios y la intención de traer aire fresco y renovado a la investigación en curso.

En este trabajo voy a dirigir mi atención a las respuestas electrofisiológicas (PEs y oscilaciones neuronales) previamente estudiadas por la vigente literatura, usando nuevos paradigmas experimentales y introduciendo nuevos conceptos a múltiple niveles con tal de buscar respuestas a cuestiones que en todos estos años de investigación han permanecido difíciles de abordar u olvidadas. Las preguntas que a continuación propongo pueden extender y profundizar nuestro conocimiento actual sobre la flexibilidad de los mecanismos cerebrales de monitorización de errores y el inicio de un nuevo camino en este prometedor campo de investigación.

1. ¿Hasta qué punto el sistema monitorización de errores en el mPFC es sensible a eventos conflictivos no-atendidos que son propensos a causar errores y que no pueden ser accedidos de forma consciente?
2. ¿De qué forma el entrenamiento durante el aprendizaje de una nueva habilidad motora conlleva a cambios funcionales en el cerebro relacionados con la monitorización automática de errores y su evaluación consciente? ¿Cómo cambia nuestra evaluación subjetiva de los errores que cometemos

mientras aprendemos una nueva tarea? ¿Y cómo cambian nuestros niveles de control cognitivo a lo largo de ese proceso?

3. ¿Cómo nos reconocemos como los agentes de nuestras acciones? ¿Será que nuestro cerebro utiliza las mismas redes neuronales para monitorizar los errores que cometemos y para evaluar la sensación de controlar nuestras acciones cuando por ejemplo perdemos el control sobre ellas? Imaginemos errores que no son intencionalmente generados por nosotros pero causados por un agente externo '*alienígeno*'?
4. ¿Será que todos los seres humanos tienen la misma tolerancia a los errores? El sentido común nos dice que no. Entonces, como estas diferencias pueden explicar distintos procesos de toma de decisión en contextos complejos que implican la monitorización de varias fuentes de información conflictiva propensas a causar errores ¿En qué medida estas diferencias individuales son reflejadas por mecanismos neurofisiológicos corticales que indexan estados cognitivos y motivacionales específicos subyacentes al monitoreo de información negativa, conflictiva o redundante?

## 9.2 Objectivos

Esta tesis incluye cuatro estudios dirigidos a explorar distintos aspectos de los mecanismos electrofisiológicos subyacentes a la monitorización de errores en diferentes contextos.

En la presente investigación se han diseñado nuevos paradigmas experimentales así como reciclado tareas experimentales ya conocidas (por ejemplo, la tarea Flanker) en los se ha modificado el enfoque inicial aportando un punto de vista diferente y creativo en pos de contestar a preguntas que han permanecido intangibles y desconsideradas durante los años previos de investigación, pero que son cruciales para entender la flexibilidad de los sistemas cerebrales implicados en la detección de errores de distinta naturaleza. En esta tesis, se han combinado novedosos paradigmas con medidas electrofisiológicas, PE y el análisis de tiempo-frecuencia de la actividad eléctrica cerebral, para examinar diferentes señales neuronales relacionadas con el error en varios contextos con la intención de aumentar la validez ecológica de estas respuestas neuronales. Los objetivos de cada uno de los estudios desarrollados en esta tesis se presentan de forma concisa a continuación.

En el **estudio 1**, hemos explorado la posibilidad de que la monitorización y la compensación de eventos no-atendidos (i.e. procesados fuera de nuestro foco consciente atencional), pero que son conflictivos y propensos a causar errores, son mediados por el sistema de errores del mPFC, lo que tradicionalmente ha sido asociado a procesos atencionales y conscientes. En este estudio se registraron PEs mientras los participantes realizaban una nueva variante de la tarea de Erikson Flanker, en la cual se manipuló la atención espacial evitando el acceso consciente a estímulos visuales que se presentan en lugares sin vigilancia (no-atendidos). La manipulación de la atención espacial de los participantes en esta tarea permitió así la exploración de índices de comportamiento y marcadores electrofisiológicos de la actividad del mPFC relacionadas con la monitorización de eventos conflictivos atendidos y no-atendidos (inconscientes). El objetivo fue investigar (i) la rapidez del mPFC en detectar dichos eventos conflictivos no-atendidos y en implementar ajustes de compensación, y (ii) estudiar qué mecanismos neurales oscilatorios constituyen la base de estos procesos. Para lograr este objetivo hemos combinado técnicas de PEs, análisis de tiempo-frecuencia y análisis de localización de la fuente de actividad eléctrica neuronal.

En el **estudio 2** intentamos examinar cómo los índices electrofisiológicos (PEs) relacionados con la monitorización de errores y su evaluación consciente cambian durante la adquisición de nuevas habilidades motoras. Para lograr este objetivo se ha diseñó un nuevo experimento en el que participantes sin formación musical aprendieron a reproducir una serie de patrones rítmicos musicales a lo largo de 12 ensayos de aprendizaje, mediante la sincronización de movimientos de *tapping* con eventos auditivos que presentaban diferentes intervalos de tiempo (rítmicos). Luego se examinaron los cambios en señales de PEs en distintas etapas de aprendizaje con el fin de entender cómo la práctica conduce a cambios funcionales en los mecanismos cerebrales asociados a la monitorización de errores y su evaluación consciente.

En el **estudio 3** hemos investigado como nuestra sensación de agencia, i.e. la sensación de ser responsables de controlar nuestras acciones, es afectada por errores experimentados en nuestro cuerpo que no son intencionalmente causados por nosotros, pero generados por un agente externo. La idea es examinar si la supervisión de errores autogenerados es computada por las mismas redes cerebrales que supervisan nuestra sensación de estar en control de nuestras acciones y evalúan, por este motivo, violaciones en nuestra experiencia natural de agencia. ¿Seré el sistema de monitorización de acción del mPFC responsable de la evaluación de estos dos tipos de errores?

En este estudio hemos aprovechado los actuales avances tecnológicos en realidad virtual para estudiar respuestas electrofisiológicas (PEs) en participantes encarnados (*embodied*) en un cuerpo virtual (un *avatar*) mientras realizaban una tarea de tiempo de reacción propicia para causar errores en un entorno de realidad virtual. Es importante resaltar que en ciertas ocasiones los participantes fueron engañados respecto a sus propias acciones, experimentando violaciones en su sensación de agencia. Concretamente, en dichas situaciones experimentales, los individuos se enfrentaban a una experiencia anormal en la que observaban cómo movimientos correctamente iniciados por ellos eran falsificados por movimientos "erróneos" de su cuerpo virtual, violando sus intenciones internas y provocando un error ajeno impuesto por el avatar. Esta configuración nos permitió la observación de diferentes señales de PEs en respuesta a errores autogenerados y a errores no intencionados (violaciones de agencia) causados por el agente virtual o ‘alienígena’.

En el **estudio 4**, hemos centrado nuestra atención en un nuevo concepto, el de *tolerancia al error*: la tendencia de los seres humanos en seguir planes de acción pre-programados (p. ej.: rutinas bien conocidas) en presencia de información contextual conflictiva, redundante o negativa, sin la necesidad de cambiar drásticamente cursos pre-seleccionados de acción. En este estudio hemos explorado en qué medida las bases neurofisiológicas de la tolerancia al error exponen aspectos cognitivos y motivacionales que subyacen a diferentes estrategias adoptadas por los seres humanos para hacer frente a información negativa del ambiente en entornos de gran incertidumbre que implican tomas de decisión compleja y que son propensos a generar errores de planificación. Para explorar diferencias individuales respecto a la *tolerancia al error,* se observó el comportamiento de 80 participantes sanos en una tarea de aprendizaje probabilístico que implicaba la toma de decisiones basadas en reglas aprendidas y luego se comparó respuestas de PEs de dos grupos de participantes (con alta y baja tolerancia al error) que difieren drásticamente en su propensión a mantener reglas recién aprendida después de recibir *feedbacks* negativos o conflictivos.

Esperamos encontrar diferencias entre los 2 grupos en relación a respuestas de PEs relacionados con el impacto motivacional y atencional durante la monitorización de feedbacks negativos y redundantes y también diferentes tipos de errores de planificación durante el proceso de toma decisiones.

## 9.3 Resultados y Discusión

Nuestra interacción con el ambiente requiere de sistemas cerebrales que evalúen el progreso del comportamiento en función de nuestros objetivos y consecuentemente, que detecten y corrijan nuestros errores de una forma adaptativa y flexible.

El estudio de las bases neurofisiológicas asociadas a la monitorización de los errores humanos ha avanzado enormemente en las últimas dos décadas. Gran cantidad de estudios han tenido una influencia importante en la caracterización de las redes cerebrales y señales electrofisiológicas asociadas a la comisión de errores, detección de conflictos y el procesamiento de eventos inesperados o sorprendentes que requieren la implementación de conductas de regulación y adaptación. Este campo ha florecido y ha permitido el desarrollo de una línia de investigación muy interesante relacionada con los procesos de control cognitivo.

En la presente tesis he intentado contribuir a este prometedor campo de investigación abordando nuevas preguntas y aportando nuevas ideas para estudiar los mecanismos neuralesimplicados en la monitorización de los errores en contextos más extensos y de mayor validez ecológica en los cuales los seres humanos interactúan, así como sus dinámicas temporales. La intención principal fue proporcionar un marco amplio e inclusivo de cara a entender varios aspectos relacionados con la naturaleza de los errores y su evaluación a nivel de los sistemas cerebrales.

Creo que las cuestiones que han sido abordadas a lo largo de esta Tesis son de importancia relevante para comprender el alcance y la profundidad con la que los mecanismos cerebrales de supervisión de errores se pueden explorar desde una perspectiva más ecológica.

### 9.3.1 La monitorización y regulación de eventos conflictivos no-atendidos es mediada por la corteza medial prefrontal

En el primer estudio de esta tesis hemos examinado la evolución temporal de señales neuronales implicadas en la monitorización de eventos conflictivos no-atendidos (no conscientes) propensos a causar errores. El estudio se llevó a cabo mediante el registro de PEs durante la realización de una nueva versión de la tarea Flanker por parte de los participantes, en la cual se manipuló la atención espacial para obstaculizar el acceso consciente de estos eventos conflictivos no-atendidos. Nuestro diseño experimental demostró que estímulos no-atendidos, que no pueden ser evaluados de manera consciente ni reportados de forma precisa (Lamme et al., 2003), pueden influenciar nuestro rendimiento, afectando a las tasas de error e influenciando la velocidad de nuestras acciones para garantizar la inhibición de respuestas erróneas automáticamente preactivadas por este tipo de estímulos.

Hemos observado que la monitorización de eventos conflictivos bien atendidos así como de no-atendidos es moderada por el mPFC. Este tipo de eventos desencadenan actividad eléctrica negativa en zonas medio-frontales de la corteza que están coordenadas por oscilaciones en theta como lo observado durante la comisión de errores (Cavanagh et al., 2009; Nigburg et al., 2008; Pastotter et al., 2013; Marco-Pallares et al., 2008). El uso de métodos de localización de las fuentes generadoras de esta actividad eléctrica confirmó que la actividad oscilatoria en theta observada durante la comisión de errores y durante la monitorización de conflictos atendidos y no-atendidos propensos a causar errores se originó en el ACC, una región del mPFC crucial en la monitorización de nuestras acciones y procesos de control cognitivo (Botvinick et al., 2001; Holroyd &Coles, 2004; Ridderinkhof et al., 2004). Un resultado importante de esta investigación fue que estas respuestas medial-frontales de actividad oscilatoria en theta  predicen el grado en el que los participantes ajustan la velocidad de sus acciones en respuesta a eventos conflictivos tanto atendidos como no-atendidos, lo que sugiere un papel de la actividad oscilatoria theta durante procesos de control inhibitorio reactivos que son desencadenados automáticamente para suprimir tendencias de respuesta erróneas incluso en la ausencia de procesos de control atencional *top-down* (Jackson & Shallice, 2001; Posner et al., 1990; Posner & DiGirolamo, 1998).

Estos resultados están de acuerdo con investigaciones recientes que muestran la influencia de las oscilaciones medial-frontales theta durante la implementación de los procesos de control cognitivo en una variedad de contextos que implican la supervisión de errores/conflictos y su regulación (Cavanagh et al, 2009; Narayanan et al, 2014; Pastotter et al, 2013). Además, este estudio amplía aún más la investigación anterior sobre el papel del mPFC en el control cognitivo, demostrando que esta región no sólo evalúa nuestras acciones sinó que también tiene un papel activo en su regulación.

En resumen, la relevancia de este trabajo recao en la desmostración de que una cantidad sustancial de información visual que se procesa fuera del foco de nuestra atención y que no está disponible al acceso consciente, activa mecanismos de la corteza medial prefrontal previamente asociada con  procesos de control *top-down*. Asimismo sugiere que la actividad oscilatoria en theta de la corteza medial prefrontal es responsable de la supervisión y regulación de acciones potencialmente inapropiadas que se activan automáticamente, incluso por los eventos ambientales a los que permanecemos no-conscientes, extendiendo hallazgos actuales sobre el papel de estas oscilaciones neuronales en el control flexible de nuestra conducta.

### 9.3.2 Procesos de monitorización de errores y su evaluación consciente durante el aprendizaje de nuevas habilidades motoras

Disponemos de una gran cantidad de evidencias sobre los mecanismos cerebrales involucrados en el procesamiento de errores; sin embargo, se ha investigado menos sobre el rol de estos mecanismos durante la adquisición de nuevas habilidades motoras. Con el fin de indagar más en esta última cuestión, en el estudio 2 (capítulo 4) hemos inspeccionado cómo la adquisición de un nuevo *skill* motor (p. ej.: aprender ritmos musicales) conlleva a cambios en la actividad eléctrica cerebral relacionada con la detección de errores y su evaluación consciente. Para lograr este objetivo, se registró la actividad de EEG de individuos sin formación musical mientras aprendían una serie de patrones rítmicos musicales. La habilidad de aprender patrones rítmicos musicales fue acompañada de cambios en señales de PEs comúnmente asociadas con la monitorización automática de errores/conflictos y los procesos de detección consciente de dichos errores.

Se observó una gran actividad de mecanismos de monitorización de errores del mPFC, reflejada por una mayor actividad del componente ERN en las primeras etapas del aprendizaje de ritmos. Este resultado demuestra que durante las fases iniciales de aprendizaje, cuando la ejecución de las tareas es más difícil, el sistema de monitorización del mPFC es crucial para el seguimiento de un gran número de tendencias de respuesta inadecuados o conflictivas, y para señalar la necesidad de aumentar el nivel de control atencional con el fin de mejorar el rendimiento (Botvinick et al, 2001.; Carter et al, 1998; Nachev et al, 2006;. Yeung et al, 2004). Estos resultados están de acuerdo con otros estudios de neuroimagen que muestran el papel destacado de redes medial-frontales durante las primeras etapas de aprendizaje de habilidades motoras que requieren de la implementación de los procesos de control *top-down* (Jenkins et al, 1994;. Jueptner et al, 1997;. Ramnani y Passingham, 2001; Toni et al, 1998).

Por otra parte, se observó que la actividad electrofisiológica asociada al procesamiento consciente de errores solo emergió durante las etapas posteriores del aprendizaje, cuando las representaciones de los ritmos se volvieron más sólidas e internalizadas y, en consecuencia, los errores se hicieron más evidentes al intérprete. A nivel neurofisiológico, la identificación consciente de errores estuvo asociada a actividad cortical positiva del EEG en zonas centroparietales de la corteza. Estos resultados apuntan a una disociación entre la monitorización de los errores y su evaluación consciente (véase Endrass et al, 2005;. Overbeek et al, 2005; Steinhauser y Yeung, 2010). Sorprendentemente, esta actividad positiva centroparietal del EEG se vio incrementada en los participantes que reproducían los patrones rítmicos con un mayor nivel de competencia. Estudios anteriores han propuesto que patrones de actividad neuronal similares a los observados en este estudio reflejan la acumulación de evidencias internas que subyacen a la consciencia de acciones o eventos motivacionalmente relevantes y la consecuente decisión sobre la exactitud de una acción dada (Steinhauser y Yeung, 2010; Yeung y Summerfield, 2012 ). Estos hallazgos sugieren que durante la adquisición de nuevas habilidades la evaluación consciente de errores, o el proceso metacognitivo que conduce a la experiencia consciente del error (Yeung y Summerfield, 2012), es contingente al estado de aprendizaje del intérprete y, en consecuencia, depende de la consolidación de las representaciones internas de la tarea.

Esta investigación extendió el trabajo previo sobre los mecanismos de procesamiento de errores y control cognitivo en tareas de tiempo-reacción a contextos con más validez ecológica que imitan muchas situaciones cuotidianas en las que aprendemos nuevas habilidades.

### 9.3.3 Errores autogenerados y errores causados por violaciones de agencia son evaluados por distintas redes neuronales

Normalmente las acciones que realizamos con nuestro cuerpo nos dan una sensación de control sobre ellas. Sin embargo, hay casos clínicos de personas que experimentan experiencias anormales en sus sentimientos de agencia y en la sensación de control sobre sus acciones, esquizofrénicos con síntomas psicóticos o pacientes con el síndrome del brazo alienígena “alien-hand”, son algunos ejemplos de estos casos (veáse p. ej.: Frith et al., 2000). Estos casos extraños y remotos nos hacen cuestionar si nuestro cerebro es capaz de distinguir acciones erróneas que son auto-generadas (p. ej.: echar sal en una taza de café como consecuencia de una distracción) de acciones erróneas que son causadas sin la voluntad del sujeto (p. ej.: mi mano agarra repentinamente la sal sin mi intención y lo vierte en la taza de café).

En el estudio 3 (capítulo 5) de esta Tesis hemos inventado una situación experimental novedosa y pionera en la que se registró la actividad de EEG de sujetos sanos incorporados (‘*embodied*’) en cuerpos virtuales (avatares) con el fin de esclarecer esta cuestión. Usando técnicas de análisis de PE hemos comparado índices neurofisiológicos relacionados con errores motores autogenerados y errores que fueran causados por el agente virtual o alienígena, es decir, sin la intención del sujeto responsable de comandar sus acciones y que perturbaron, momentáneamente, su experiencia natural de agencia.

Nuestros resultados demuestran que errores auto-generados y errores-alienígenas están asociados a distintas señales PE, con diferente actividad cortical espacial y diferentes latencias de procesamiento. Estos resultados sugieren la existencia de dos sistemas cerebrales independientes, pero no exclusivos, responsables de la supervisión de nuestras acciones y de la evaluación de nuestro sentido de control sobre esas acciones, o nuestro sentido de agencia.

Errores motores autogenerados desencadenaron la actividad del componente frontocentral ERN justo después del inicio de un movimiento erróneo (100 ms post inicio de la respuesta). Esta señal neuronal refleja el funcionamiento de un sistema interno de monitorización de errores del mPFC, que utiliza modelos directos internos del plan motor (‘efference copy’) para detectar errores motores durante el curso de nuestros movimientos (Coles et al, 2001; Gehring et al, 1993; Rodríguez-Fornells et al, 2002). Este sistema es responsable de monitorizar acciones erróneas e implementar correcciones rápidas o comportamientos compensatorios sin depender de información sensorial o propioceptiva.

Curiosamente, la inserción de errores-alienígenas en acciones correctamente ejecutadas por los participantes, lo que disminuyó de forma importante su experiencia de agencia, suscitó una señal negativa más lenta (alrededor de 400 ms) en zonas parietales de la corteza. Esta señal podría reflejar la activación de un sistema cerebral que registra la ocurrencia de violaciones de agencia o la pérdida de la sensación de control motor, y que a nivel neurofisiológico no está relacionada con redes medio-frontales del cerebro implicadas en procesos de monitorización de nuestras acciones, como monitorización de errores, conflictos o resultados negativos (véase Ullsperger et al., 2014 para una revisión de este tema). De acuerdo con algunas teorías, este sistema parece utilizar copias del comando motor para predecir (predicciones internas) si el resultado de la acción que se va a producir es coherente con el feedback externo generado por dicha acción, y en casos de que esta comparación no detecte ningún error o incongruencia, sentimos control o nos sentimos agentes de esa acción (Frith et al., 2000; Syfnozik et al., 2008). Como este sistema depende en última instancia de la tramitación de diferentes feedbacks re-aferentes que llegan a las regiones corticales somatosensorial, visual y auditiva con diferentes latencias, es probable que el tiempo requerido por el cerebro humano para detectar errores o violaciones entre las predicciones internas y las consecuencias reales del movimiento tenga un retraso de 350 a 400 ms (la latencia de la señal neural observada durante errores-alienígenas).

Además, la distribución parietal de esta señal neural tras violaciones de agencia converge con estudios previos que enfatizan la importancia de las regiones parietales posteriores en relación a la sensación disminuida de la agencia (Farrer et al., 2003; Farrer et al., 2008) y a procesos subyacentes a intenciones de movimiento (Desmurget et al, 2009; Sirigu et al, 2004).

Cabe destacar que la amplitud de la señal neurofisiológica de violaciones de la agencia se correlacionó con la sensación subjetiva de la propiedad del cuerpo virtual, lo que sugiere una relación entre las representaciones internas de la imagen corporal y el impacto que las acciones impuestas no intencionalmente por el agente alienígena ejercen en la experiencia natural fenoménica de agencia.

Curiosamente, esta respuesta neuronal de violaciones de agencia mostró una gran similitud en su topografía y latencia con otras firmas electrofisiológicas relacionadas con violaciones semánticas o conceptuales, específicamente con el componente N400 (Kutas y Federmeier, 2011). Una intrigante cuestión que surge aquí es en qué medida las violaciones de agencia son evaluadas de una manera similar a una violación semántica conceptual.

Este trabajo reveló nuevas evidencias sobre las dinámicas temporales del cerebro humano para diferenciar errores autogenerados y errores-alienígenas así como sus centros de procesamiento, proporcionando nuevas evidencias sobre los mecanismos cerebrales implicados en la integración de la información interna y sensorial en la acumulación de un sentido coherente de agencia durante la monitorización de nuestras acciones. Por otra parte, el marco experimental presentado aquí ofrece una oportunidad única para estudiar la experiencia de agencia de una forma más ecológica abriendo nuevos caminos para la investigación sobre los procesos mentales que subyacen a la atribución de agencia en poblaciones sanas y clínicas.

### 9.3.4 Evidencias electrofisiológicas de las diferencias individuales en la tolerancia al error en humanos

Los humanos difieren considerablemente en la manera cómo interpretan y evalúan los errores y el feedback informativo de un error acerca de su rendimiento en el desarrollo de una tarea. Estas diferencias resultan fundamentales para el entendimiento de la divergencia intra-individual durante la iniciación y el mantenimiento del control cognitivo, así como para poder racionalizar sobre los procesos implicados en la toma de decisiones de los humanos en entornos complejos e inciertos y en los que la probabilidad de cometer un error es muy elevada.

Por este motivo, en el último estudio de esa tesis (estudio 4, capítulo 6), investigamos si las diferencias individuales en la *tolerancia al error*, i. e. la tendencia (o *bias*) de los humanos a seguir planes de acción bien establecidos (e. g. rutinas bien conocidas) aún y la presencia de indicios conflictivos o feedbacks informando de un error, pueden estar reflejados en determinados mecanismos neurofisiológicos asociados con distintos estados cognitivos y emocionales.

Para evaluar diferencias neurofisiológicas en la tolerancia al error, examinamos los PEs de dos grupos extremos en relación a la tolerancia al error (categorizados como grupos de baja y alta tolerancia al error) que, durante la realización de una tarea de aprendizaje con la inclusión de toma de decisiones basadas en unas normas, mostraron diferencias drásticas en la propensión a mantener nuevas reglas aprendidas tras recibir tanto feedbacks engañosos como feedbacks redundantes sobre errores. Las diferencias en la tolerancia al error fueron caracterizadas por diferentes tipos de errores de planeamiento y por unos patrones específicos encontrados en la actividad eléctrica cerebral durante la anticipación y evaluación del feedback informativo del error. Los participantes con mayor tendencia a mantener las nuevas reglas aprendidas, i. e. con mayor tolerancia al error, mostraron menor compromiso atencional y motivacional durante la anticipación de los feedbacks que aparecían ensayo tras ensayo informando de su rendimiento en la tarea. Este hecho se evidenció en los datos electrofisiológicos con una reducción de la amplitud de la SPN, componente de los ERPs caracterizado por una negatividad y que es sensitivo a los estados anticipatorios atencionales y motivacionales a los feedbacks informativos (Brunia et al., 2011; Fuentemilla et al., 2013; Moris et al., 2013). Los participantes con alta tolerancia al error también mostraron una reducción en la actividad cortical durante el procesamiento del feedback, evaluado por los componentes FRN y P3 (Fischer & Ulllsperger, 2013; Yeung et al., 2005). Por consiguiente, los participantes en los que su toma de decisiones depende mayormente en un conocimiento interior basado en la regla, parecen tener menor sensibilidad y dependencia a la información proporcionada vía feedback externo, operando consecuentemente en un modo independiente al feedback. En contraste, los individuos con baja tolerancia al error mostraron una tendencia más elevada a cambiar frecuentemente su conducta tras recibir feedbacks negativos, dirigiendo y mostrando más atención a esta información externa.

Por otro lado, resulta interesante el hecho que aunque los individuos con alta y baja tolerancia al error fueron afectados de forma distinta tanto por los feedbacks engañosos o redundantes), en todos ellos, una vez habían adquirido evidencias suficientes de que las nuevas contingencias de la tarea implicaban un una reprogramación de los cursos preseleccionados de sus acciones, sus patrones neurales de actividad eléctrica cerebral (predictivos de los reajustes conductuales basados en reglas) convergían en un mismo tipo, sugiriendo unos mecanismos corticales adaptativos comunes para la actualización de la información del feedback en relación al grado de tolerancia al error.

Los resultados de este estudio respaldan la existencia de unos procesos cognitivos y motivacionales específicos que subyacen a las diferencias individuales en la tolerancia al error entre los humanos, pudiendo influenciar la omisión de diferentes tipos de acciones erróneas deliberadas durante procesos complejos de toma de decisiones.

Creemos que estos hallazgos son relevantes y enfatizan la importancia de tener en cuenta las diferencias individuales relacionadas con la supervisión de las cadenas de comando complejas (Norman, 1988), así como también contribuyen a la investigación actual centrada en la comprensión de los procesos mentales existentes tras la falibilidad humana en entornos complejos propensos a la comisión de errores.

## 9.4 Conclusión

En la presente tesis he introducido nuevas herramientas conceptuales y planteé un amplio conjunto de restricciones de varios niveles sobre las concepciones actuales de la neurociencia cognitiva sobre el constructo “error”. Asimismo, he desarrollado nuevos paradigmas experimentales para expandir nuestro entendimiento actual sobre los mecanismos cerebrales implicados en el procesamiento de errores en varios contextos en los cuales los seres humanos interaccionan.

Las conclusiones de este trabajo probablemente aportarán nuevos conocimientos sobre varios aspectos de los mecanismos cerebrales que subyacen a los procesos cognitivos relacionados con el procesamiento de errores, abriendo nuevas vías para futuras investigaciones centradas en estos fenómenos en contextos de mayor validez ecológica.

Esta obra reúne diferentes líneas de investigación que, en la mayoría de los casos, ha progresado de forma independiente al estado actual de la investigación. El presente trabajo abraza importantes tópicos de la Neurociencia Cognitiva actual, tales como: (i) el papel de la atención y de los procesos (in)conscientes en procesos superiores de control cognitivo; (ii) el papel de los mecanismos de control cognitivo y monitorización y acción durante procesos de aprendizaje motor; (iii) la experiencia de agencia y control sobre las acciones; y (iv) las diferencias individuales durante los procesos de toma de decisiones.

Basándonos en los experimentos presentados en esta Tesis podemos extraer las siguientes conclusiones:

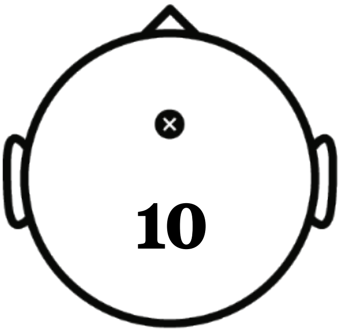
1. Hay una gran cantidad de información no-atendida en el ambiente y que es procesada de forma inconsciente la cual influye en nuestro comportamiento. Esta información, cuando es relevante, activa redes neurales de control prefrontales, tradicionalmente asociadas con los procesos de control consciente y de *top-down*. La actividad neural oscilatoria en theta parece presentarse como un mecanismo neurobiológico plausible por el cual redes medial-frontales son capaces de monitorizar y regular de forma flexible y adaptativa acciones inapropiadas que son desencadenadas de forma automática por información ambiental no-atendida.
2. La práctica conduce a cambios funcionales en los mecanismos cerebrales asociados a la monitorización de los errores y a su evaluación consciente. En las etapas tempranas del aprendizaje de una nueva habilidad motora, por ejemplo, hay un gran compromiso de los mecanismos de la corteza medio-frontal encargados de monitorizar y solventar planes motores erróneos y conflictivos y de aumentar el nivel de control atencional. La conciencia del error, a su vez, depende del estado de las representaciones internas asociadas con la tarea, que se refinan con la práctica de tal manera que la experiencia consiente del error está íntimamente relacionada con el estado de aprendizaje del intérprete. Estos dos procesos están asociados a distintos mecanismos neurofisiológicos.
3. Los errores motores autogenerados y los errores que reflejan violaciones de agencia son monitorizados por sistemas cerebrales distintos. Los errores autogenerados se basan exclusivamente en un sistema de control rápido interno que opera en la corteza medial prefrontal sobre mecanismos predictivos de control motor. Los errores que reflejan violaciones de

agencia, a su vez, requieren la participación de un sistema de control más lento y que opera en las regiones parietales del cerebro. La experiencia fenoménica de control de nuestras acciones, o de agencia, depende de representaciones semánticas previas de nuestra imagen corporal (predicciones sensoriales).

4. Las diferencias individuales relativas a la tolerancia al error se caracterizan por distintos tipos de error de planteamiento y distintos patrones de actividad cortical durante la anticipación y evaluación de la información externa negativa, redundante o conflictiva. Estas diferencias exponen estados cognitivos y motivacionales que diferencian la personalidad de varios agentes en contextos de gran incertidumbre propensos a la comisión de varios tipos de errores. Estos resultados subrayan la importancia de toener en cuenta las diferencias individuales relacionadas con el procesamiento de errores en contextos que implican complejos procesos de toma de decisión.

Creo que estos hallazgos y las ideas presentes a lo largo de esta tesis ofrecen una gran contribución a la investigación actual en la neurociencia cognitiva sobre la monitorización y el control adaptativo de nuestro comportamiento.

Estas evidencias tienen gran relevancia para validar las teorías actuales sobre los procesos neurofisiológicos de procesamiento de errores y los mecanismos de control cognitivo asociados. Fundamentalmente, este trabajo ofrece nuevas perspectivas para comprender el alcance y la profundidad a la que el sistema de control de errores en los seres humanos puede ser estudiado en contextos más ecológicos que imitan las complejas demandas cognitivas de la vida cotidiana.



# 10

# References

# REFERENCES

Adams, J. A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior, 3,* 111-150.

Aglioti, S., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nat Neurosci, 11,* 1109-1116.

Albert, R., Jeong, H., & Barabasi, A.-L. (2000). Error and atack tolerance of complex networks. *Nature, 406,* 378-382.

Allain, S., Burle, B., Hasbroucq, T., & Vidal, F. (2009). Sequential adjustments before and aftter partial errors. *Psychonomic Bulletin & Review, 16,* 356-362.

Allport, D. A., Tipper, S. P., & Chmiel, N. R. J. (1985). Perceptual integration and postcategorical filtering. In M.I.Posner & O. S. M. Marin (Eds.), *Mechanisms of attention: Attention &: performance XI* ( Hillsdale, NJ: Erlbaum.

Angel, R. W. & Higgins, J. R. (1969). Correction of false moves in pursuit tracking. *J.Exp.Psychology, 82,* 185-187.

Arbel, Y. & Donchin, E. (2009). Parsing the componential structure of post-error ERPs: a principal component analysis of ERPs following errors. *Psychophysiology, 46,* 1179-1189.

Baldwin, K. B. & Kutas, M. (1997). An ERP analysis of implicit structured sequence learning. *Psychophisiology, 34,* 74-86.

Bandura, A. (1989). Self-regulation of motivation and action through internal standards and goal systems. In L.A.Pervin (Ed.), *Goal concepts in personality and social psychology* (pp. 19-85). Hillsdale, NJ: Erlbaum.

Barcelo, F., Escera, C., Corral, M. J., & Perianez, J. A. (2006). Task switching and novelty processing activate a common neural network for cognitive control. *J.Cogn Neurosci., 18,* 1734-1748.

Barnes, T., Kubota, Y., Hu, D., Jin, D. Z., & Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature, 437.*

Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nat Neurosci, 10,* 1214-1221.

Bernstein, P. S., Scheffers, M. K., & Coles, M. G. H. (1995). Where did I go wrong?' A psychophysiological analysis of error detection. *J.Exp.Psychol.: Human Perception and Performance, 21,* 1312-1322.

Bonini, F., Burle, B., Liegeois-Chauvel, C., Regis, J., Chauvel, P., & Vidal, F. (2014). Action monitoring and medial frontal cortex: leading role of supplementary motor area. *Science, 343,* 888-891.

Botvinick, M. & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature, 391,* 756.

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

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

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn Sci., 16,* 106-113.

Broadbent, D. E. (1958). *Perception and Communication*. London: Pergamon.

Broadbent, D. E. (1982). Task combination and selective inIake of information. *Acta Psychologica, 50,* 253-290.

Brown, R. M., Chen, J. L., Hollinger, A., Penhune, V. B., Palmer, C., & Zatorre, R. J. (2013). Repetition Suppression in Auditory-Motor Regions to Pitch and Temporal Structure in Music. *J.Cogn.Neurosci., 25,* 313-328.

Brunia, C. H., Hackley, S. A., van Boxtel, G. J., Kotani, Y., & Ohgami, Y. (2011). Waiting to perceive: reward or punishment? *Clin.Neurophysiol., 122,* 858-868.

Burle, B., Roger, C., Allain, S., Vidal, F., & Hasbroucq, T. (2008). Error negativity does not reflect conflict: A reappraisal of conflict monitoring and anterior cingulate cortex activity. *Journal of Cognitive Neuroscience, 20,* 1637-1655.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerey, S. C., & Rauch, S. L. (1998). The counting Stroop: An interference task specialized for functional neuroimaging-validation study with functional MRI. *Human Brain Mapping, 6,* 270-282.

Buzsaki, G. & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science, 304,* 1926-1929.

Camara, E., Rodriguez-Fornells, A., & Munte, T. F. (2009). Functional Connectivity of Reward processing in the brain. *frontiers in human neuroscience, 1,* 1-14.

Camara, E., Rodriguez-Fornells, A., & Munte, T. F. (2010). Microstructural brain differences predict functional hemodynamic responses in a reward processing task. *J Neuroscience, 30,* 11398-11402.

Cañas, J. J., Quesada, J. F., Antoli, A., & Fajardo, I. (2003). Cognitive flexibility and adaptability to environmental changes in dynamic complex problem-solving tasks. *Ergonomics, 46,* 482-501.

Carter, C. S. & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cogn Affect.Behav.Neurosci., 7,* 367-379.

Carter, C., Braver, T., Barch, D., Botvinick, M., Noll, D., & Cohen, J. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280,* 747-749.

Casey, S. M. (2006). *The atomic chef: and other true tales of design, technology, and human error*. Santa Barbara, CA: Aegean Pub. Co.

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *J Neuroscience, 29,* 98-105.

Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaption in reinforcement learning. *Neuroimage., 49,* 3198-3209.

Cavanagh, J. F., Zambrano-Vasquez, L., & Allen, J. J. B. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology, 49,* 220-238.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008a). Listening to musical rhythms recruits motor regions of the brain. *Cereb.Cortex, 18,* 2844-2854.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008b). Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *J.Cogn.Neurosci., 20,* 226-239.

Cohen, J. D., Botvinick, M. M., & Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: who's in control? *Nat Neurosci, 3,* 421-423.

Cohen, M. A., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in Cognitive Sciences, 16,* 411-417.

Coles, M. G. H., Gratton, G., Bashore, T. R., Ericksen, C. W., & Donchin, E. (1985). A psychophysiological investigation of the continuous flow model of human information processing. *J.Exp.Psychol.: Human Perception and Performance, 11,* 529-553.

Coles, M. G. H., Scheffers, M. K., & Holroyd, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biological Psychology, 56,* 189.

Danielmeier, C., Eichele, T., Forstmann, B., Tittgemeyer, M., & Ullsperger, M. (2011). Posterior medial frontal cortex activity predicts post-error adaptations in task-related visual and motor areas. *J.Neurosci., 31,* 1780-1789.

Danielmeier, C. & Ullsperger, M. (2011). Post-error adjustments. *Front Psychology, 2,* 233.

Daprati, E., Franck, N., Georgieff, N., Proust, J., Pacherie, E., Dalery, J. et al. (1997). Looking for the agent: an investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition, 65,* 71-86.

Debener, S., Ullsperger, M., Slegel, M., Flehler, K., von Cromon, Y. D., & Engel, A. K. (2005). Trial-by-Trial Coupling of Concurrent Electroencephalogramand Functional Magnetic Resonance Imaging Identifies theDynamics of Performance Monitoring. *J Neuroscience, 25,* 11730-11737.

Dehaene, S. & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition, 79,* 1-37.

Delorme, A. & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J Neuroscience Methods, 134,* 9-21.

Desmurget, M. & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences, 4,* 423-431.

Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., & Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science, 324,* 811-813.

Donchin, E. & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11,* 357-374.

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

Eichele, T., Debener, S., Calhoun, V. D., Specht, K., Engel, A. K., von Cramon, Y. D. et al. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *PNAS, 105,* 6173-6178.

Endrass, T., Franke, C., & Kathmann, N. (2005). Error awarenes in a saccade countermanding task. *J Psychophysiol, 19,* 275-280.

Eriksen, B. A. & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16,* 143-149.

Falkeinstein, M., Hohnstein, J., Hoormann, J., & Blanke, L. (1990). Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In C.H.M.Brunia, A. W. K. Gaillard, & A. Kok (Eds.), *Psychophysiological brain research* (pp. 192-195). Tilburg.

Falkeinstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology, 51,* 87-107.

Farrer, C., Bouchereau, M., Jeannerod, M., & Franck, N. (2008). Effect of distorted visual feedback on the sense of agency. *Behav.Neurol., 19,* 53-57.

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

Ferdinand, N. K., Mecklinger, A., & Kray, J. (2008). Error and Deviance Processing in Implicit and Explicit Sequence Learning. *J.Cogn.Neurosci., 20,* 629-642.

Fields, R. E., Wright, P. C., & Harrison, M. D. (1995). A Task Centered Approach to Analysing Human Error Tolerance Requirements. In (pp. 18-26).

Finney, S. A. & Palmer, C. (2003). Auditory feedback and memory for music performance: sound evidence for an encoding effect . *Mem.Cognition, 31,* 51-64.

Fischer, A. G. & Ullsperger, M. (2013). Real and Fictive Outcomes Are Processed Differently but Converge on a Common Adaptive Mechanism. *Neuron, 79,* 1243-1255.

Floyer-Lea, A. & Matthews, P. M. (2005). Distinguishable brain activation networks for short- and long-term motor skill learning. *J Neurophysiol, 94,* 512-518.

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.

Foti, D. & Hajcak, G. (2009). Depression and reduced sensitivity to non-rewards versus rewards: evidence from event-related potentials. *Biol.Psychol, 81,* 1-8.

Fries, P. (2005). Mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences, 10,* 474-480.

Frith, C. (2005). The self in action: lessons from delusions of control. *Conscious.Cogn, 14,* 752-770.

Fuentemilla, L., Cucurell, D., Marco-Pallares, J., Guitart-Masip, M., Moris, J., & Rodriguez-Fornells, A. (2013). Electrophysiological correlates of anticipating improbable but desired events. *Neuroimage., 78,* 135-144.

Gabriel, M. (2002). Cingulate Cortex. In V.S.Ramachandran (Ed.), *Enciclopedia of the Human Brain* (pp. 775-791).

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

Gehring, W. J. & Fencisik, D. E. (2001). Functions of the medial frontal cortex in the processing of conflict and errors. *J Neuroscience, 21,* 9430-9437.

Gehring, W. J., Gross, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychol Sci, 4,* 385-390.

Gehring, W. J. & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science, 295,* 2279-2282.

Gemba, H., Sasaki, K., & Brooks, V. B. (1986). "Error" potentials in lim-bic cortex (anterior cingulate area 24) of monkeys during motor learning . *Neurosci Lett, 70,* 223-227.

Gratton, G., Coles, M. G. H., Sirevaag, E. J., Eriksen, C., & Donchin, E. (1988). Pre- and poststimulus activation of response channels: a psychophysiological analysis. *J.Exp.Psychol.Hum.Perform., 14,* 331-344.

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *J.Exp.Psychology: General.*

Gruendler, T. O., Ullsperger, M., & Huster, R. J. (2011). Event-related potential correlates of performance-monitoring in a lateralized time-estimation task. *PLoS One, 6.*

Haggard, P. & Chambon, V. (2012). Sense of Agency. *Curr.Biol., 22,* 390-392.

Hester, R., Foxe, J. J., Molholm, S., Shpaner, M., & Garavan, H. (2005). Neural mechanisms involved in error processing: a comparison of errors made with and without awareness. *Neuroimage, 27,* 602-608.

Hillyard, S. A. & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *PNAS, 95,* 781-787.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical Signs of Selective Attention in the Human Brain. *Science, 182,* 177-180.

Hillyard, S. A., Squires, K. C., Bauer, J. W., & Lindsay, P. H. (1971). Evoked potential correlates of auditory signal detection. *Science, 172,* 1357-1360.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Phil.Trans.R.Soc.Lond.B, 353,* 1257-1270.

Hochman, E., Orr, J., & Gehring, W. (2014). Toward a more sophisticated response representation in theories of medial frontal performance monitoring: The effects of motor similarity and motor asymmetries. *Cereb.Cortex, 24,* 414-425.

Holroyd, C. B. & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol.Rev., 109,* 679-709.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G. H. et al. (2004). Dorsal anterior cinculate cortex shows fMRI respone to internal and external error signals. *Nat Neurosci, 7,* 497-498.

Holroyd, C. B., Nieuwenhuis, S., Mars, R. B., & Coles, M. G. H. (2004). Anterior cingulate cortex, selection for action, and error processing. In M.I.Posner (Ed.), *Cognitive neuroscience of attention* (pp. 219-231). New Yourk: Guilford Press.

Holroyd, C. B., Yeung, N., Coles, M. G., & Cohen, J. D. (2005). A mechanism for error detection in speeded response time tasks. *J.Exp.Psychol Gen., 134,* 163-191.

Houde, J. F. & Nagarajan, S. S. (2011). Speech production as state feedback control. *Front.Hum.Neurosci., 5,* 10.3389/fnhum.2011.00082.

Husain, M. & Rorden, C. (2003). Non-spatially lateralized mechanisms in Hemispatial Neglet. *Nature Reviews Neurosci, 4,* 36.

Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science, 302,* 120-122.

Jack, A. L. & Shallice, T. (2001). Introspective physicalism as an approach to the science of consciousness. *Cognition, 79,* 161-196.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S. J., & Passingham, R. E. (1994). Motor Sequence Learning: A Study with Positron Emission Tomography. *J Neuroscience, 14,* 3775-3790.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z., & Barabasi, A.-L. (2001). The large-scale organization of metabolic network. *Nature, 407,* 654.

Jocham, G., Neumann, J., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2009). Adaptive Coding of Action Values in the Human Rostral Cingulate Zone. *J.Neurosci., 29,* 7489-7496.

Jocham, G. & Ullsperger, M. (2009). Neuropharmacology of performance monitoring. *Neurosci.Biobehav.Rev., 33,* 48-60.

Jongsma, M., Meeuwissen, E., Vos, P., & Maes, R. (2007). Rhythm perception: Speeding up or slowing down affects different subcomponents of the ERP P3 complex. *Biol.Psychology, 75,* 219-228.

Jueptner, M., Stephan, K. M., Frith, C. D., Brooks, D. J., Frackowiak, R. S. J., & Passingham, R. E. (1997). Anatomy of motor learning. I. Frontal cortex and attention to action. *J.Neurophysiol., 77,* 1313-1324.

Jung, T., Makeig, S., Westerfield, M., Townsend, J., Courchene, E., & Sejnowski, T. J. (2001). Analysis and Visualization of Single-Trial Event-Related Potentials. *Human Brain Mapping, 14,* 166-185.

Kawashima, S., Ueki, Y., Kato, T., Matsukawa, N., Mima, T., Hallet, M. et al. (2012). Changes in striatal dopamine release associated with human motor-skill acquisition. *PLoS One, 7,* e31728.

Kerns, JG., Cohen, J., MacDonald, A., Cho, R., Stenger, V., & Carter, C. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303,* 1023-1026.

King, J., Korb, F., von Cromon, Y. D., & Ullsperger, M. (2010). Post-error behavioral adjustments are facilitated by activation and suppression of task-relevant and task-irrelevant information processing. *J.Neurosci., 30,* 12759-12769.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J.Neurosci., 21,* RC159.

Kotani, Y., Kishida, S., Hiraku, S., Suda, K., Ishii, M., & Aihara, Y. (2003). Effects of information and reward on stimulus-preceding negativity prior to feedback stimuli. *Psychophysiology, 40,* 818-826.

Krigolson, O. E. & Holroyd, C. B. (2006). Evidence for hierarchical error processing in the human brain. *Neuroscience, 137,* 13-27.

Kutas, M. & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu.Rev.Psychol, 62,* 621-647.

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review, 111,* 880-913.

Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences, 7,* 12-18.

Lange, K. (2011). The reduced N1 to self-generated tones: an effect of temporal predictability? *Psychophysiology, 48,* 1088-1095.

Lau, H. C. & Passingham, R. E. (2007). Unconscious Activation of the Cognitive Control System in the Human Prefrontal Cortex. *J Neuroscience, 27,* 5805-5811.

Levelt, W. (1989). *Speaking: from Intention to articulation.* Cambridge,MA: MIT Press.

Longo, M. R., Schuur, F., Kammers, M. P., Tsakiris, M., & Haggard, P. (2008). What is embodiment? A psychometric approach. *Cognition, 107,* 978-998.

Luu, P., Collins, P., & Tucker, D. (2000). Mood, personality, and self-monitoring: Negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *Journal of Experimental Psychology: General, 129,* 43-60.

Luu, P., Tucker, D. M., & Makeig, D. (2004). Frontal midline theta and the error-related negativity: neurophysiological mechanisms of action regulation. *Clinical Neurophysiology, 115,* 1821-1835.

Luu, P., Tucker, D. M., Derryberry, D., Reed, M., & Poulsen, C. (2003). Electrophysiological responses to errors and feedback in the process of action regulation. *Psychological Science, 14,* 47-53.

Luu, P. & Tucker, D. (2001). Regulating Action: Alternative activation of midline frontal and motor cortical networks. *Clinical Neurophysiology, 112,* 1295-1306.

Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2010). Nobody Is Perfect: ERP Effects Prior to Performance Errors in Musicians Indicate Fast Monitoring Processes. *PLoS.One., 4.*

Mangun, G. R., Hillyard, S. A., & Luck, S. J. (1993). Electrophysiological substracts of visual selective attention. In D.E.Meyer & S. Kornblum (Eds.), *Attention and Performance XIV: Sinergies in Experimental Psychology, Artificial Inteligence, and Cognitive Neuroscience* ( Cambridge, MA: MIT Press.

Marco-Pallares, J., Camara, E., Munte, T. F., & Rodriguez-Fornells, A. (2008). Neural mechanisms underlying adaptive actions after slips. *J.Cognitive Neuroscience, 20,* 1595-1610.

Mars, R. B., Coles, M. G. H., Grol, M., Holroyd, C. B., Nieuwenhuis, S., Hulstijn, W. et al. (2005). Neural dynamics of error processing in medial frontal cortex. *Neuroimage, 28,* 1007-1013.

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience, 10,* 647-656.

Miller, E. K. & Cohen, M. X. (2001). An integrative theory of prefrontal cortex function. *Annu Rev Neurosci., 24,* 167-202.

Miller, J. (1991). The flanker compatibility effect as a function of visual angle, attentional focus, visual transients, and perceptual load: A search for boundary conditions. *Perception & Psychophysics, 49,* 270-288.

Miltner, W., Braun, C., & Coles, M. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a generic neural system for error detection. *J.Cognitive Neuroscience, 9,* 788-798.

Moris, J., Luque, D., & Rodriguez-Fornells, A. (2013). Learning-induced modulations of the stimulus-preceding negativity. *Psychophysiology, 50,* 931-939.

Muller, S., Moller, J., Rodriguez-Fornells, A., & Munte, T. (2005). Brain potentials related to self-generated and external information used for performance monitoring. *Clinical Neurophysiology, 116,* 63-74.

Murphy, P. R., Robertson, I. H., Allen, D., Hester, R., & O'Connell, R. G. (2012). An electrophysiological signal that precisely tracks the emergence of error awareness. *Front.Hum.Neurosci., 6,* 1-16.

Nachev, P., Wydell, H., O'Neill, K., Hussain, M., & Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *Neuroimage, 36,* 155-163.

Nager, W., Kholmetz, C., Altenmuller, E., Rodriguez-Fornells, A., & Munte, T. F. (2003). The fate of sounds in conductors' brains: an ERP study. *Cogn.Brain Res., 17,* 83-93.

Nakamura, K., Sakai, K., & Hikosaka, O. (1998). Neuronal Activity in medial frontal cortex during learning of sequential procedures. *J Neurophysiol, 80,* 2671-2687.

Narayanan, N., Cavanagh, J. F., Frank, M. J., & Laubach, M. (2013). Common medial frontal mechanisms of adaptive control in humans and rodents. *Nat Neurosci, 16,* 1888-1895.

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience and Biobehavioral Reviews, 28,* 441-448.

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: evidence from an antisaccade task. *Psychophysiology, 38,* 752-760.

Nieuwenhuis, S., Yeung, N., Holroyd, C., Schurger, A., & Cohen, J. (2004). Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. *Cereb.Cortex, 14,* 741-747.

Nigburg, R., Cohen, M. X., Ridderinkhof, K. R., & Sturmer, B. (2012). Theta Dynamics Reveal Domain-specific Control over Stimulus and Response Conflict. *J.Cogn Neurosci., 24,* 1264-1274.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review, 88,* 1-15.

Norman, D. A. (1988). *The Design of Everyday Things.* New York: Doubleday.

Norman, D. A. & Shallice, T. (1986). Attention to action: willed and automatic control of behavior. In R.J.Davidson, D. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* ( New York: Plenum.

O'Connell, R. G., Dockree, P. M., Bellgrove, M. A., Kelly, S. P., Hester, R., Garavan, H. et al. (2007). The role of cingulate cortex in the detection of errors with and without awareness: a high-density electrical mapping study. *Eur.J.Neurosci., 25,* 2571-2579.

O'Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat.Neuroscience, 15,* 1729-1735.

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., & Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20 s before they occur. *J.Neurosci., 29,* 8604-8611.

Oliveira, F. T., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J.Cognitive Neuroscience, 19,* 1994-2004.

Overbeek, T. J. M., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable Components of Error Processing. *J Psychophysiol, 19,* 319-329.

Padrao, G., Mallorqui, A., Cucurell, D., Marco-Pallares, J., & Rodriguez-Fornells, A. (2013). Neurophysiological differences in reward processing in anhedonics. *Cogn Affect.Beha v.Neurosci., 13,* 102-115.

Palmer, C. & Drake, C. (1997). Monitoring and planning capacities in the acquisition of music performance skills. *Ca n J Exp Psychol, 51,* 369-384.

Pastotter, B., Dreisbach, G., & Bauml, K. H. (2013). Dynamic Adjustments of Cognitive Control: Oscillatory Correlates of the Conflict Adaptation Effect. *J.Cogn Neurosci., 25,* 2167-2178.

Paulus, M. P. & Frank, L. R. (2006). Learning the value of information in an uncertain world. *Neuroimage, 15,* 836-846.

Paus, T. (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Na t Rev Neurosci, 2,* 417-424.

Paus, T., Koski, L., Caramanos, Z., & Westbury, C. (1998). Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex. *NeuroReport, 9,* 37-47.

Pavone, E. F., Marzi, C. A., & Girelli, M. (2009). Does subliminal visual perception have an error-monitoring system? *Eur J Neurosci., 30,* 1424-1431.

Petersen, S. E., van Mier, H., Fiez, J. A., & Raichle, M. E. (1998). The effects of practice on the functional anatomy of task performance. *PNAS, 95,* 853-860.

Pfordresher, P. Q. (2006). Coordination of perception and action in music performance . *Adv.Cogn.Psychol., 2,* 183-198.

Pfordresher, P. Q. & Palmer, C. (2006). Effects of hearing the past, present, or future during music performance. *Perception & Psychophysics, 68,* 362-376.

Pfordresher, P. Q., Palmer, C., & Jungers, M. K. (2007). Speed, accuracy, and serial order in sequence production. *Cogn Sci., 31,* 63-98.

Picton, T. W. (1992). The P300 wave of the human event-related potential. *J.Clin.Neurophysio., 9,* 456-479.

Picton, T. W., Lins, O. G., & Scherg, M. (1995). The recording and analysis of event-related potentials. In F.Boller & J. Grafman (Eds.), *Ha ndbook of Neuropsychology* (pp. 3-73). Amsterdam: Elsevier.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin.Neurophysiol., 118,* 2128-2148.

Posner, M. I. (1994). Attention: the mechanisms of consciousness. *PNAS, 91,* 7398-7403.

Posner, M. I. & DiGirolamo, G. L. (1998). Executive attention: Conflict, target detection and cognitive control. In R.Parasuraman (Ed.), *The Attentive Bra in* (pp. 401-423). Cambridge, MA: MIT Press.

Praamstra, P., Turgeon, M., Hesse, C. W., Wing, A. M., & Perryer, L. (2003). Neurophysiological correlates of error correction in sensorimotor-synchronization. *Neuroimage, 20,* 1283-1297.

Procyk, E., Tanaka, Y. L., & Joseph, J. P. (2000). Anterior cingulate activity during routine and non-routine sequential behavious in macaques. *Nat.Neuroscience, 3,* 502-508.

Rabbitt, P. (1966a). Error correction time without external error signals. *Nature,* 212-238.

Rabbitt, P. (1966b). Errors and error correction in choicer esponse tasks. *Journal of Experimental Psychology, 71,* 264-272.

Rabbitt, P. & Rodgers, B. (1977). What does a man do after he makes an error? an analysis of response programming. *Q.J.Exp.Psychol, 29,* 727-743.

Ramnani, N. & Passingham, R. E. (2001). Changes in the Human Brain during Rhythm Learning. *J.Cogn.Neurosci., 13,* 952-966.

Reason, J. (1990). *Human Error.* Cambridge University Press.

Ridderinkhof, K. R., Ramautar, J. R., & Wijnen, J. G. (2009). To P(E) or not to P(E): a P3-like ERP component reflecting the processing of response errors. *Psychophysiology, 46,* 531-538.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306,* 443-447.

Rodriguez-Fornells, A., Kurzbuch, A. R., & Munte, T. F. (2002). Time course of error detection and correction in humans: neurophysiological evidence. *J.Neurosci., 22,* 9990-9996.

Rougier, N. P., Noelle, D. C., Braver, T. S., Cohen, J. D., & O'Relly, R. C. (2005). Prefrontal cortex and flexible cognitive control: Rules without symbols. *PNAS, 12,* 7338-7343.

Ruiz, M. H., Jabusch, H. C., & Altenmuller, E. (2009). Detecting Wrong Notes in Advance: Neuronal Correlates of Error Monitoring in Pianists. *Cereb.Cortex, 19,* 2625-2639.

Rushworth, M. F., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in Cogn Sci, 8,* 417.

Scheffers, M. K. & Coles, M. G. H. (2000). Performance Monitoring in a Confusing World: Error-Related Brain Activity, Judgments of Response Accuracy, and Types of Errors. *J.Exp.Psychol.: Human Perception and Performance, 26,* 141-151.

Schultz, W., Dayan, P., & Montague, P. (1997). A neural substrate of prediction and reward. *Science, 275,* 1593-1599.

Schultz, W. (2007). Behavioral Dopamine Signals. *trends in neuroscience, 30,* 203-210.

Senders, J. W. & Moray, N. P. (1991). *Human Error: Cause, Prediction, and Reduction.* Hilllsdale, NJ: Earlbaum.

Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat Rev Neurosci, 12,* 154-167.

Sheth, S. A., Mian, M. K., Patel, S. R., Asaad, W. F., Williams, Z. M., Dougherty, D. D. et al. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature, 488,* 218-221.

Shibasaki, H. & Hallet, L. (2006). What is the Bereitschafts potential? *Clin.Neurophysiol., 117,* 2341-2356.

Shima, K. & Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. *Science, 282,* 1335-1338.

Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A. et al. (2004). Altered awareness of voluntary action after damage to the parietal cortex. *Nat.Neurosci., 7,* 80-84.

Solman, G. J. F., Cheyne, J. A., & Smilek, D. (2012). Found and missed: Failing to recognize a search target despite moving it. *Cognition, 123,* 100-118.

Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Andersib, D. K. (1988). Cognitive Flexibility Theory: Advanced Knowledge Acquisition in Ill-Structured Domains. In V.Patel (Ed.), *Proceedings of the 10th Annual Conference of the Cognitive Science Society* ( Hillsdale, NJ: Erlbaum.

Steinhauser, M. & Yeung, N. (2010). Decision Processes in Human Performance Monitoring. *J Neuroscience, 30,* 15643-15653.

Sternad, D., Abe, M. O., Hu, X., & Muller, H. (2011). Neuromotor noise, error tolerance and velocity-dependent costs in skilled performance. *PLoS Comput Biol, 7.*

Sumner, P., Nachev, P., Morris, P., Petersm AM, Jackson, S., Kennard, C. et al. (2007). Human Medial Frontal Cortex Mediates Unconscious Inhibition of Voluntary Action. *Neuron, 54,* 697-711.

Sutton, R. S. & Barto, A. G. (1998). *Reinforcement learning: An introduction.* Cambridge, MA: MIT Press.

Synofzik, M., Thier, P., Leube, D. T., Schlotterbeck, P., & Lindner, A. (2010). Misattributions of agency in schizophrenia are based on imprecise predictions about the sensory consequences of one's actions. *Brain, 133,* 262-271.

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Conscious.Cogn, 17,* 219-239.

Taylor, P. J. C., Nobre, A. C., & Rushworth, M. F. (2007). Subsecond changes in top-down control exerted by the human medial frontal cortex during conflict and action selection: a combined transcranial magnetic stimulation-electroencephalography study. *J.Neurosci., 27,* 11343-11353.

Toni, I., Krams, M., Turner, R., & Passingham, R. E. (1998). The Time Course of Changes during Motor Sequence Learning: A Whole-Brain fMRI Study. *Neuroimage, 8,* 50-61.

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiol.Rev., 94,* 35-79.

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends Cogn Sci, 18,* 259-267.

van Driel, J., Ridderinkhof, K. R., & Cohen, M. X. (2012). Not all errors are alike: theta and alpha EEG dynamics relate to differences in error-processing dynamics. *J Neuroscience, 32,* 16795-16806.

van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, S., & Lamme, V. A. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *J Neuroscience, 28,* 8053-8062.

Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nat Neurosci, 2,* 229-238.

Walton, M. E., Devlin, J. T., & Rushworth, M. F. (2004). Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience, 7,* 1259-1265.

Wiegmann, D. A. & Shappell, S. A. (2003). *A human error approach to aviation accident analysis.* Aldershot, UK: Ashgate.

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D. et al. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *PNAS, 90,* 8722-8726.

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nat.Rev.Neurosci., 12,* 739-751.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science, 269,* 1880-1882.

Wolpert, D. M. & Miall, R. C. (1996). Forward Models for Physiological Motor Control. *Neural Netw., 9,* 1265-1267.

Woods, D. D. (1994). *Behind Human Error: Cognitive Systems, Computers, and Hindsight.* Ohio: Crew System Ergonomics Information Analysis Center.

Woods, D. L., Hillyard, S. A., Courchesne, E., & Galambos, R. (1980). Electrophysiological signs of split-second decision-making. *Science, 207,* 655-657.

Yacubian, J., Glascher, J., Schroeder, K., Sommer, T., Braus, D. F., & Buchel, C. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *Journal of Neuroscience, 26,* 9530-9537.

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP Correlates of Feedback and Reward Processing in the Presence and Absence of Response Choice. *Cereb.Cortex, 15,* 535-544.

Yeung, N. & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philos.Trans.R.Soc.Lond B Biol.Sci., 367,* 1310-1321.

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,* 959.

Yin, H., Mulcare, S. P., Hilario, M., Clouse, E., Holloway, T., Davis, M. et al. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat Neurosci, 12,* 333-341.

Understanding the nature of the brain mechanisms involved in the flexible evaluation of human action and the adaptive changes that follow behavioral imperceptions, or errors, is a basic goal of modern cognitive neuroscience. The past 20 years have been prolific regarding this topic. Still crucial challenges for cognitive neuroscientists remain; as for example the development of novel methods and paradigms that allow the study of the neural mechanisms of error-monitoring, and its temporal dynamics, in broad ecological contexts that reproduce the complexity of everyday life situations in which human performance is likely to break down.

The present dissertation provides new alternatives and contributions regarding this issue. This work addresses novel questions, applies new toolkits and brings new ideas in order explore various aspects of the neurophysiology underlying error-monitoring and cognitive control processes in more extended and ecological contexts in which humans interact.