TESIS DOCTORAL

Programa de Doctorado en Psicología

Neural Mechanisms of Preparation in Expectation and Selective Attention

Doctorando

José María González Peñalver

Directores

María Ruz Cámara
Carlos González García

UNIVERSIDAD
DE GRANADA

Departamento de Psicología Experimental

Diciembre de 2023

Editor: Universidad de Granada. Tesis Doctorales
Autor: José María González Peñalver
ISBN: 978-84-1195-193-7
URI: https://hdl.handle.net/10481/89809

If dreams can't come true,
Then why not pretend?

— The Frog (Over the Garden Wall, 2014)

A Tania.

Agradecimientos

Llevo pensando en los agradecimientos desde mucho antes de empezar a escribir la tesis. No porque fuera tan tranquilo como para encontrarme en mitad de mi propio cuento de la lechera (aunque la leche un poco sí que me la llevé). Era más bien el ancla a lo que es posible. Si puedo pensar en los agradecimientos -me decía- es porque en el fondo creo que puedo escribir la tesis. Y bien, aquí estamos. La tesis está escrita. Saldrá mejor, o peor, pero *existe*. Y los agradecimientos, en consecuencia, van a existir también. No toda la gente agradecida los va a leer, ni toda la gente que los lea será agradecida. Me disculpo de antemano con ambas partes. Yo os quiero igual.

Gracias, María, por enseñarme lo emocionante que es enfrentarse a cosas que son difíciles, sí, pero que al final se pueden. Ya sea un paper incomprensible, ya sea una tesis (espero) algo más comprensible. Mil veces lo he dicho y al menos una más lo diré: nunca habría hecho un doctorado si no nos hubiéramos cruzado en clase de neurociencia. No habría investigado, no habría una tesis, y viviría desconociendo que, sin ciencia, la realidad es mucho menos interesante. Como se ya se ha comentado, vaya donde vaya ahora (siempre en compañía de mis dramas), me alegro infinitamente de que hayamos dado el uno con el otro. Gracias.

Gracias, Carlos, por sacar mil huecos cada vez que me atranco *continuamente.* Por ser el mejor ejemplo de a qué aspirar en este mundillo. Y bueno, gracias a tu padre, por asegurar que el linaje González vaya dirigiendo todas las etapas de mi vida. Emocionado pensando en qué me van a dirigir tus hijos, si los tuvieras.

Bueno, vamos aligerando que me enrollo y al final de estos agradecimientos hay un libro bastante chulo. Os animo a que si os quedáis con ganas de más le echéis un ojo.

Mil millones de gracias al equipo conocido en mi casa como *esta gente*. Todo el mundo tiene su *esta gente.* Tengo mucha suerte de que vosotros seáis la mía. Gracias: a JC, por acordarte de mí más que yo mismo. Por ser la persona con más talento para enseñar que conozco, ya sea a conducir o a tirar un balón. Por ser manquísimo en el Mario Kart. Porque, sin ti, no existiría nada parecido a *esta gente.* A Josu, por entenderme. Por pasar a preguntar cómo voy. Por dejarme saber cómo vas tú. Por haber hecho que siempre que estuviéramos en tu piso, fuera cual fuera, me sintiera como en casa. A Rafa, por tener más paciencia conmigo de la que me he ganado. Por creer que sé más de lo que sé, y en el camino, darme razones para seguir aprendiendo. Por apilar las piedras de "buen camino", para que mientras te sigo sepa que no voy a perderme. A Greta por, literalmente, salvarme la vida. Enciendo fuegos en mitad de la lluvia. Enseñándome a hacer arroz *en una sartén que ni es la tuya* (flipo). Porque sé que cuando me cuentas lo de que tienes una amiga que trabaja en Disneyland lo dices porque me vas a invitar a que vayamos de un momento a otro. Y porque, cuando me lleves (que me llevarás) lo menos importante del viaje será el destino. A Andrés, por rebosar amor en todo lo que haces y todo lo que dices. Por abrazarme y quererme aun cuando sabes que voy completamente recubierto de pelos de la gata. Hay que ser muy valiente para querer tan bien.

A veces *esta gente* se expande. Y entra quien siempre me falta, y a quien siempre encuentro. Gracias: a Klara, por tirar por tierra mi teoría de que "tus mejores amigos los conoces como tarde en la carrera". Por construir conmigo una de mis relaciones favoritas. Por cómo hemos pasado de ser un dúo cómico natural, a un festival de lágrimas, risas y abrazos que me devuelven años de vida. Porque vaya agosto nos hemos pegado, *la virgen*. A Luis, por respetar el más que evidente potencial de una conversación. Por aceptar mis repetidos intentos, a cada cual menos exitoso, de transmitirte lo significativa que es mi relación contigo.

A las personas *de aquí*, que están desde hace mil años. Que igual por Granada ya si eso nos vemos en Navidad, pero que siguen estando *aquí.* A Eloy, por recordarme que siempre hay alternativas. Porque siempre se puede hacer una videollamada. Porque siempre, siempre, estás ahí. A Carlos, por transmitirme que vale la pena intentarlo. Por sacarme de mi casa, aun cuando vivo completamente apalancado. Por seguir respetando mis gustos cinematográficos, aunque todavía no haya ni una puñetera película que nos guste a los dos. A Anna, por acogerme igual de bien en cualquier país. Por Marvel, obvio. Porque en adelante, Eurovisión sin ti va a tener cero gracia. A Victoria, porque en serio, te quiero un montón. Me encanta verte y me encanta sentir que no ha pasado el tiempo cada vez que te veo.

Y como la ciencia no se hace sola, hay un rinconcillo en el CIMCYC donde se reúne gente tremendamente apañada. El 345 ha sido, literalmente, mi segunda casa. Gracias: a Ana Paqui, por cuidar de mi desde que vine en grado creyéndome el más listo del barrio. Por conversaciones hiper dramáticas en la playa. Por saber cuándo preguntarme cómo estoy. A Paloma, por ser, a lo tonto, la persona del despacho con la que he compartido más intereses (not including science). Por guiarme durante la tesis. Por animarme a cantar, saliera como saliera aquello. A Alberto, por darme el espacio seguro para el drama que necesitaba. Por los trapicheos con los jueguicos. Por ser un ejemplo de que elegir no es renunciar. A Blanca, por enseñarme lo poco que realmente sé sobre el mundo. Por ser, durante mucho tiempo, la balsa que mantenía el 345 a flote. A Marta, 345 breve pero fundamental, por salvarme la vida con el experimento de reso. Cero metáforas aquí. A Jose, por tener más motivación en cada una de las 40000 cosas que haces de la que cualquiera tendría en toda una vida. Pero en serio, trabaja un poco menos muchacho. A Paco y Paula, por resucitar el 345. Por devolverme las ganas de subir al CIMCYC. Porque ojalá haber sido yo la mitad de competente de lo que sois vosotros. Gracias por todo lo que vais a traer.

Aunque hay otros números que uno también asocia a buena ciencia, y mejor gente. A la gente del 343, gracias: A Luis Ciria, el post-doc más prometedor del CIMCYC. Por apadrinar a cada persona que entra al doctorado. Por aligerar la carga. A Chiara, porque la ciencia es importante, pero el salseo lo es mucho más. Por ser insoportablemente buena en todo lo que haces, y compartir parte de ello conmigo. A Clara, porque sigo pensando que si me pego un poco más igual se me pega parte de esa energía.

Mucha gente me ha apoyado durante estos años. Pero es absolutamente imposible haber llegado hasta aquí sin la continua labor de salvación que han desplegado los ingenieros. Gracias en general al grupo de Juanma y Javi por llevarme de erasmus de disciplina científica. Y en concreto, gracias a: Juan Eloy, porque en serio detengan a esta belleza. Por ser tan absurdamente amable y, simplemente, bueno, que fuiste capaz de anclar nuestra amistad a mis exabruptos de mala sombra, y encima funcionó. A Carmen, por ser el mayor descubrimiento que aporta esta tesis. Por buscarme para un proyectillo que al final acabó siendo mucho más que un par de análisis. Por salvarme la vida más veces que páginas hay aquí escritas. A David, por enseñarme el 90% de las cosas que aprendido durante el doctorado. Por ser mi primer compañero, y dejar el listón bien alto.

Obviamente, a la gente del CIMCYC. Gracias: a Omar, por abrir constantemente espacios donde expresarse, ya sean agobios, inseguridades, frikadas o delirios. A Conchi, por cuidarnos a mí y a mis latentes ganas de bailar. A Cris, por hacerme sentir la persona más graciosa del planeta, justo después de ti. A María Cobos, por ser la persona con la que tengo el mayor ratio intensidad/conversación. No es que hablemos todos los días, pero cuando toca, toca. A Javi, por contarme una tarde en tu despacho lo que era el decoding. Ya te devolveré el favor. A Isma, porque siempre se necesita un abrazo y recibir algo de feedback positivo por, absolutamente, cualquier cosa. A Carmen, por ser hipnóticamente interesante, por ser increíblemente amable. A Tao, por las excursiones pendientes. Por interesarte por mí, y por todos mis compañeros. A Marta, por acompañarme en este camino en paralelo. Por sincronizar los agobios y las celebraciones. Por pasar de compi, a amiga. A Sofía, por seguir aquí casi 10 años después. Por ese café que tenemos pendiente. A Alba, por ese viaje al caribe que nos vamos a pegar cuando abramos el tarro. A Giorgia, por las cervezas al sol. Y también a Belén, Fer, Ana, Dani, Jeanette, Nuria, Pablo Rodríguez, Melody, Filip y Mariagrazia, por aguantar mis dramas, mis intentos de ser gracioso, por compartir las mil horas que llevamos, y al menos las mil más que nos quedan.

Hay un par de personas muy importantes que no sabía en qué apartado poner, así que vamos a poner “gente guay de la carrera”. Por ejemplo. Gracias a: María (selectividad), porque si no hubieras adivinado que nos iba a caer Nietzsche en selectividad igual no estamos aquí hoy. Por infinitas casualidades que me hacen dudar de que lo sean. Por seguir ahí. A Laura, por los mejores puñeteros abrazos. Patenta eso ya, en serio. Por la comedia.

Hay otro grupillo sumamente importante (más, probablemente, que todos los demás). Aunque no creo que vayan a leer esto, esta es mi tesis y yo le escribo a quien me da la gana. Y no me podían haber ayudado ni un poco más. Gracias: al Bicho, por los duros combates. Por no dejar ni un pantalón sin babas. A Nala, por acompañarme en aquellas carreras pandémicas y esperarme al final de cada vuelta. Os echo de menos, a los dos. A Pumba, porque no es que seas un pesao, es que nos quieres mucho. A Shadow, por seguirme con la mirada mientras me voy (más que tu madre, qué fuerte). Por las peleillas matutinas. Tú entrena que ya verás como algún día ganas. A Sarita, que tendría que estar aquí conmigo. A Sally, lo más parecido al líquido que puede ser un vertebrado. Y, sobre todo, a mis dos reinas. A Arya, por atravesarme con la mirada cada vez que abres los ojos. Porque te he visto más veces panza arriba que de pie. Y a Ágata, por ser la mejor decisión que cualquiera podría tomar. Por estar dormida en mi cuello mientras escribo estas palabras. Por aquella vez que me borraste tres páginas de la tesis y casi me da un infarto.

Y finalmente, a mi familia. Gracias a: mis padres, por preguntarme todos los días que cómo me iba el cole, y cómo se portaba mi seño. Por llevarme, traerme, cuidarme, y aguantarme. Porque todavía os va a tocar aguantar dos horas de defensa en inglés. A mi hermana Carmen, por cuidar del piso y de Ágata siempre que lo necesito. Ya que estás, apúntate la contraseña de Netflix. A la abuela Lola, por recordarme que lo primero es antes, y que yo soy un valiente. A mi tía Carmen, por ser prácticamente mi tercera abuela. Por plantar el interés por la enseñanza.

Gracias, Adriana. Realmente, podría escribir dos o tres tesis sobre las razones que tengo para darte las gracias. Pero vamos a coger un par para darle envidia a quien se lea esto, y el resto te las cuento luego en persona. Gracias, por haberme acompañado durante el 100% de esta etapa, desde que empecé el máster hasta hoy. Has conocido a una versión de mí bastante particular, y has decidido que eso bastaba. Gracias por saber cuándo toca esforzarse y cuando toca dejarse llevar. Gracias, porque, en serio, no me puedo creer la suerte que tengo.

Gracias, gracias, Tania. Por siempre saber lo que necesitaba antes que yo mismo. Por creer en mi mucho más que yo mismo. Porque, sin ti, esto no es lo mismo. Por las videollamadas furtivas, por la determinación ante cualquier reto. Por compartir con todos una parte de tu fuerza. Gracias por esta tesis.

Contents

Figures and Tables

Acronyms

ACC	Anterior Cingulate Cortex
acc	Accuracy (behavioral)
AI	Anterior Insula
Att	Attention
AUC	Area Under de Curve
BOLD	Blood-Oxygen-Level-Dependent
COBIDAS	Committee on Practices in Data Analysis and Sharing
CTI	Cue-Target Interval
CTP	Canonical Template Pattern
CTT	Canonical Template Tracking
DLPFC	Dorsolateral Prefrontal Cortex
EEG	Electroencephalography
ERP	Event-Related Potential
Exp	Expectation
FDR	False Discovery Rate
FEF	Frontal Eye Fields
fMRI	functional Magnetic Resonance Imaging
FO	Frontal Operculum
FP	Frontal Pole
FPN	Frontoparietal Network
FWE	Family Wise Error
GLM	General Linear Model
HC	Hippocampus
HRF	Hemodynamic Response Function

ICA	Independent Component Analysis
IFS	Inferior Frontal Sulcus
IOG	Inferior Occipital Gyrus
IPC	Inferior Parietal Cortex
IPS	Intraparietal Sulcus
IS	Intraparietal Sulcus
IT	Inferotemporal cortex
ITG	Inferior Temporal Gyrus
LDA	Linear Discriminant Contrast
LDC	Linear Discriminant Contrast
LiG	Lingual Gyrus
LOC	Lateral Occipital Cortex
LOSO	Leave-one-subject-out
MD	Multiple Demand Network
MEG	Magnetoencephalography
MEG	Magnetoencephalography
MFC	Medial Frontal Cortex
MFG	Middle Frontal Gyrus
MVPA	Multivariate Pattern Analysis
OcG	Occipital Gyri
OCP	Occipital Pole
PANGEA	Power ANalysis for GEneral ANOVA designs
pCC	Posterior Cingulate Cortex
PMd	Dorsal Premotor Cortex
PRC	Precentral Gyrus

RDM	Representational Dissimilarity Matrix
RLPFC	Rostrolateral Prefrontal Cortex
ROC	Receiver Operative Curve
ROI	Region of Interest
RSA	Representational Similarity Analysis
RTs	Reaction Times
S1	Primary Somatosensory Cortex
SMA	Supplementary Motor Area
SMG	Supramarginal Gyrus
SPC	Superior Parietal Cortex
SPL	Superior Parietal Lobule
STC	Superior Temporal Cortex
STG	Superior Temporal Gyrus
SVM	Support Vector Machine
TGM	Temporal Generalization Matrix
TPJ	Temporoparietal Junction
V1	Primary visual cortex
VFC	Ventral Frontal Cortex
VVC	Visual Ventral Cortex
WM	Working Memory

Chapter I

INTRODUCTION

The world is complex and ever-changing. We are bombarded by large numbers of stimuli that generate ambiguous patterns in our sensory receptors, to which we usually need to give fast and effective responses. To face this complexity and ambiguity, among other things humans use the capacity to anticipate what is about to happen before it does. Indeed, it is widely assumed that preparing for incoming information improves performance (Barton et al., 2006). However, some fundamental questions arise from this seemingly evident statement. How does preparation improve performance? Does it entail the representation of specific incoming information? And, crucially, is proactive preparation a unified process, or does it recruit dissociable mechanisms in different cognitive contexts?

To exemplify this last question, we can describe two scenarios. For the first one, let us think about someone who has been to a movie theatre. When the movie is over and they are ready to leave, they find out that it is heavily raining. With no umbrella or solution in sight, they decide to take a taxi to get home. It is then logical to think that, before they see a taxi, they will actively look for it and that in doing so, the image, idea, or representation of a taxi could presumably take place in their brains. This will help them detect a taxi as soon as it enters their visual field. Now, in a different situation, this same person has taken a plane, and has just landed at the airport. In this case, they do not actually need a taxi, since they already booked the bus ticket to get to the city. However, when they leave the airport, they see a long line of tractors waiting in line to pick up passengers, right where taxis are supposed to be. They would undoubtedly be surprised. But why? Although it was not connected to their current goals, as they approached the exit of the airport, they had an image, idea, or representation of what should be lining up at the door. They were, again, prepared to see taxis. But in mechanistic cognitive terms, was the representation of the taxi equal when it was relevant for their goals than when it was just more likely to appear, but behaviorally irrelevant? This thesis is composed of three studies in which we approached this question by applying multiple analysis techniques to Electroencephalography (EEG) and functional

Magnetic Resonance Imaging (fMRI). To do so, we directly contrasted the anticipation of category-specific information in two fundamental cognitive contexts: selective attention (information relevance) and perceptual expectations (information probability).

In this thesis preparation is framed as a crucial top-down phenomenon. Notably, influential models have proposed diverse mechanisms through which anticipatory patterns interact with stimulus inputs to guide perception. Our work is structured around two of the most relevant ones: the Biased Competition model (Desimone & Duncan, 1995) and the Predictive Processing framework (Ficco et al., 2021; Friston, 2005). Nevertheless, these models currently do not address the interrelation between different preparatory phenomena or whether they share common or distinct underlying top-down mechanisms. The following pages contextualize the theoretical and empirical background that supports the research presented in this thesis. We begin by describing the relevance of endogenous brain processing and several of its manifestations, such as working memory and cognitive control. Then, the focus shifts to preparation as form of top-down processing, highlighting the nature of anticipatory representations. Finally, attention and expectation are introduced as two phenomena often conflated, which are crucial to understand the nature of proactive anticipation.

1.1. The proactive brain: control, attention and working memory

The notion that there are two main forms in which the brain processes information is not new. Early notions of this dichotomy distinguished between automatic and conscious processing (Posner & Snyder, 1975). The idea was that automatic processes are stimulus-driven, and are enacted as a consequence of perception, while conscious processes are governed by internal goals. However, more frequent accounts do not stablish the distinction based on automaticity, since automatic responses can also arise from internal states (de Lange et al., 2018). The specific distinction usually depends on the field of study. In cognitive control research they are frequently labeled reactive vs. proactive (Braver, 2012), highlighting the agency of individuals in that type of tasks. In attention, it is common to find an endogenous vs. exogenous distinction (Fernández et al., 2022), which focuses on the source of attentional orienting, either the stimulus or internal goals. Last, within the predictive processing framework, such distinction is often termed bottom-up vs. top-down (Gordon et al., 2019), which highlights the hierarchical nature of the cerebral cortex and the direction of information processing. All these definitions embrace ground rules about brain function that make these dichotomies plausible. First, there are different pathways for each form of processing, which originate in different regions. Second, the brain is organized hierarchically, such that lower levels of this hierarchy process stimulus information with high detail, while higher levels process with higher abstraction and integration (Rauss & Pourtois, 2013). And third, the two flows of information processing interact and influence each other. Crucially, proactive/endogenous processing, as indicated before, can happen in at least two manners. Although proactive cognition can be thought of a system that is primarily based on goals, internal knowledge of the statistical regularities of the world can also influence decision making and preparation, which has been extensively explored by predictive coding accounts (Summerfield & Egner, 2009). In the following paragraphs, we briefly summarize crucial findings in proactive processing, focusing on cognitive control, attention and working memory, and how they set the bases to study anticipation. We will go back to predictive processing at the end of this introduction, and directly contrast it with selective attention.

The conceptualization of cognitive control was one of the first attempts to highlight top-down processing as a mechanism that allows the implementation of specific cognitive programs guided by internal goals (Posner & Snyder, 1975). Currently, the term defines the set of mechanisms that guide behavior and thought in accordance to goals (Braver, 2012). Specifically, cognitive control, also referred to as executive functions or executive control, has been proposed to encompass three main functions:

shifting, inhibition and updating (Diamond, 2013; Miyake et al., 2000). Shifting refers to the ability to dynamically switch tasks, and was one the main functions ascribed to cognitive control in earlier models (Norman & Shallice, 1986), whereas inhibition is related to avoiding "dominant, automatic, or prepotent responses when necessary" (Miyake et al., 2000, p. 57). Finally, updating involves two crucial and interrelated functions: working memory (WM) and selective attention (Diamond, 2013). A robust finding in cognitive control is the existence of a wide network of regions that are recruited in contexts when control is needed. It is known as the Multiple Demand Network (MD, Duncan, 2010), and has been proposed to be split into two main sub-networks (Dosenbach et al., 2007; Palenciano et al, 2019). The frontoparietal network (FPN) includes the dorsolateral prefrontal cortex (DLPFC) and the Intraparietal Sulcus (IPS), and would be related to transient task switching and within-trial adjustments, being hence more related to reactive control (conflict and error processing), while the other involves the dorsal anterior cingulate cortex (dACC), the frontal operculum and rostrolateral prefrontal cortex (RLPFC). This last network, known as the cingulo-opercular network, has been related to the anticipation and tonic maintenance of different task demands and rules in cue-target paradigms in which cues give information about changes in task rules (González-García et al., 2016).

As stated above, attention can be considered a function that serves cognitive control, specially tuned to select relevant information in a system with limited capacity (Mackie et al., 2013). Although William James famously stated that "everyone knows what attention is" (James, 1890), the reality is that there are many different interpretations of the actual implications of the term (Hommel et al., 2019). Here, we consider attention as "the processes that allow an individual to select and focus on particular input for further processing while simultaneously suppressing irrelevant or distracting information" (Stevens & Bavelier, 2011, p. 30). A well-known theoretical model of attention proposes that it is composed of three specific networks: alerting, executive control, and orienting (Posner & Petersen, 1990). The alerting network sets (phasic alerting) and maintains (tonic alerting or vigilance) a general state of activation of the cognitive system, supported by right lateralized broad regions such as the thalamus and frontal and parietal cortices. The executive control network deals with interference at the moment of target processing, and is theoretically similar to proactive and reactive cognitive control, including similar regions to the MD (Braver, 2012; Dosenbach et al., 2008; Petersen & Posner, 2012). The orienting network supports the ability to select task-relevant information. It is divided into two separate networks: the ventral one consists of the temporo-parietal junction (TPJ) and the ventral frontal cortex (VFC), and has been linked to exogenous attentional capture. The dorsal attention network, on the other hand, consists of the IPS, superior parietal lobule (SPL) and the frontal eye fields (FEF), and has consistently been linked to endogenous attention. Importantly, endogenous orienting (or selection) of attention has been related to top-down connections between the FEF and the primary visual cortex (V1) in anticipation of eye movements, thus influencing perceptual processing (Pooresmaeili et al., 2014). Moreover, the orienting network has been proposed to work by biasing perceptual areas towards relevant stimuli while suppressing irrelevant information, after target onset (ventral network) and, crucially, before (dorsal network, Mackie et al., 2013). Noteworthy, this mechanism is akin to the predictions made by the influential Biased Competition model of selective attention (Desimone & Duncan, 1995), which we describe with more detail below.

In all of the crucial functions that can be carried in an endogenous manner, an ability to maintain, flexibly organize, and prioritize information, either task related rules or relevant features or spatial locations is implicitly assumed. This ability is called working memory (WM), and has been consistently related to the FPN, particularly to the prefrontal cortex a posterior parietal cortex (D'Esposito & Postle, 2015; Nir-Cohen et al., 2020) as well as perceptual regions (Ranganath et al., 2004). Areas of the FPN show increased activity when the contents of working memory are maintained or updated, regardless of the specific object (Roth et al., 2006), and it has been found that its performance can be modulated by cognitive control, as it acts as a gating mechanism that allows relevant information to enter WM (Boag et al., 2021). Overall, regions of the FPN can represent task rules, goals, or object categories. Moreover, early work found that visual regions of the occipitotemporal cortex linked to the representation of specific object categories, such as the so-called fusiform face area and parahipocampal place area, are active while maintaining face and house information in WM (Ranganath et al., 2004). This finding has been supported by work using multivariate analyses, that have shown that maintained information can be decoded from a range of stimulus-specific regions (e.g. Lee et al., 2013; Sreenivasan et al., 2014; see D'Esposito & Postle, 2015 for a detailed review). Taking all of this into account, other perspectives of the bases of WM have proposed that it does not necessarily rely on a specific network, but is better described as the result of the interaction between control, attention and perceptual regions (Bledowski et al., 2010).

Altogether, these findings suggest that proactive cognition is the result of a set of varied but highly interconnected mechanisms. A possible mechanistic description of a situation where proactive control is needed could unfold as follows: cognitive control represents task rules and allocates resources to the attentional networks that will be needed to achieve a specific goal, alertness prepares certain systems to give a fast response when a target is shown, while the orienting system of attention (or selective attention) represents the stimuli that will be relevant by maintaining its features active in perceptual regions through WM. These features can flexibly change thanks to adaptative gating systems exerted by cognitive control through the FPN. An outstanding question, however, is what sort of cognitive and neural architecture can support these top-down computations without suffering from a recursive homunculus argument, a topic that we will go back to in the discussion. The main take out from this general picture is our ability to hold information that is relevant to internally generated goals in the absence of direct perceptual stimulation, to give efficient responses. We can prepare, and we do so based on our goals or prior knowledge.

In the following section we will lay the focus on proactive preparation, why is it important, what paradigms have been used to measure it, and what models have attempted to explain how it affects perception and action.

1.2. Important considerations in the study of preparation

Both proactive control and predictive coding contribute to preparatory effects, as evidenced by reports of neural preparatory activity that has been shown for a plethora of processes, including cognitive control (Baines et al., 2011; González-García et al., 2016; Hebart & Baker, 2018), attention (Battistoni et al., 2017; Kastner et al., 1999; Nobre & Serences, 2018), expectation (Aranda et al., 2010; Kok et al., 2017), or WM (Koshino et al., 2015; van Driel et al., 2017). This preparatory activity often consists on the finding that incoming relevant, probable, or cued information is represented in the brain in a delay period that frequently follows a cue (Battistoni et al., 2017). A consistent observation in cognitive neuroscience is that anticipation preactivates brain regions partially overlapping with those engaged in subsequent stimulus processing. This phenomenon transcends various cognitive domains. Selective attention cues (Battistoni et al., 2017; Nobre & Serences, 2018) preactivate specific shape patterns in visual cortex (Stokes et al., 2009), relevant regions of space (Giesbrecht et al., 2006), and object-selective (Peelen & Kastner, 2011; Soon et al., 2013) and category (Esterman & Yantis, 2010; González-García et al., 2018) perceptual regions. Relatedly, perceptual expectation cues lead to the preactivation of specific templates of oriented gabors (Kok et al., 2017), direction (Ekman et al., 2017), motor patterns (de Lange et al., 2013) or abstract shapes (Hindy et al., 2016). Although it seems to be a consistent pattern that anticipated information is represented in perceptual brain regions, before we go deeper into how information is represented in the brain, we will briefly describe our operational definition of representation in this thesis.

1.2.1. Representations

In the field of neuroscience, representations, as described by Poldrack (2021), can be thought of as the systematic relationship between characteristics of the physical internal and external world and the neural activity of the brain. In certain cases, the large-scale spatial arrangement of neural reactions aligns with the world's structure (isomorphic representations), as seen in the retinotopic organization of visual cortex responses (Fishman, 1997) or the tonotopic organization in the auditory cortex (Langers & van Dijk, 2012). In these, neighboring neurons react to neighboring regions of the corresponding stimulus space (e.g. regions that are adjacent in the visual space elicit activity in neurons that are also adjacent in V1). However, in other scenarios, like the case of place cells in the hippocampus (O'Keefe & Dostrovsky, 1971), the overarching spatial arrangement of neuronal responses does not mirror the external world. All of this implicitly assumes that every neuron has a preference for a certain stimulus or feature in a specific spatial region, the receptive field, and that preference evokes increased activity in the neuron, which in turn can be measured (e.g. firing rates, increased BOLD signal, etc.). In the field of visual perception, preferences can go from simple features such as line orientation (Hubel & Wiesel, 1959) to complex stimuli such as faces (Kanwisher et al., 1997). However, although a neuron will give the strongest response for stimuli in its receptive field, it can fire for stimuli for which it does not have a strong preference (Dumoulin & Wandell, 2008). Since popular neuroimaging methods such as fMRI or EEG cannot track activity of individual neurons, research focuses on "population receptive fields", which account for the features to which a set of neurons (usually, sampled in one or multiple voxels, in the case of fMRI) gives the strongest response (Wandell & Winawer, 2015).

When studying populations of neurons, some relevant issues come up. Let us say we have a set of different stimuli and want to study how they are represented in a certain population. A priori, we have as many explanations (dimensions) as neurons are found within that population (Badre et al., 2021). However, these dimensions can be reduced to their “representational dimensionality”, that is, the smaller number of dimensions necessary to account for the variance of such population. There are several methods that attempt at finding this minimal dimensionality, by studying the similarity of the patterns based on a priori hypotheses. In this regard, multivariate methods study the patterns of activity associated with different conditions, and allow to stablish different levels of similarity. Hence, if neurons/voxels can be grouped based on how similar their patterns of activity towards specific conditions are, we can theorize that they “represent” those dimensions. One example of these analyses is multi-voxel pattern analysis (MVPA, Norman et al., 2006), aimed at decoding different features or objects. Within this idea, if two or more conditions can be decoded from a group of voxels (that is, the pattern of voxel-wise activity is different in the chosen population for each condition), said group might be relevant in representing that dimension. Nevertheless, additional considerations and controls need to be carefully evaluated, since MVPA is sensitive to small changes in variance that may be unrelated to the experimental dimensions to classify (Hebart & Baker, 2018). This notion has challenged the aforementioned assumption of stimulus specific regions like the fusiform face area, where stimuli other than faces can also be decoded (Haxby et al., 2001). Moreover, this method is not only restricted to perceptual representations, since it has also been used to find different task sets in parietal and frontal regions (e.g. González-García et al., 2016; Hebart et al., 2018; Palenciano et al., 2019), highlighting the great variability in which representations can take place. Other forms of studying this variability go to further lengths and analyze the contribution that several representational dimensions can have on the variance of neural activity. This is the case of representational similarity analysis (RSA, Kriegeskorte et al., 2008), which measures how similar stimuli or conditions are to each other, and how this similarity reflects specific representational profiles. Crucially, representations can happen in the absence of visual stimulation, and are rooted in the same regions as actual perception. This has been shown in the context of WM (e.g. Lee et al., 2013), mental imagery (e.g. Cichy et al., 2012; Dijkstra et al., 2017) and anticipation (Battistoni et al., 2017), although this notion has also been challenged (Pace et al., 2023; Theeuwes, 2013).

1.2.2. Univariate and multivariate analyses

Throughout this introduction we mentioned several neuroimaging results that support the work presented here. However, these results come from different analysis methodologies which result in different theoretical implications, namely univariate and multivariate methodologies. Moreover, they arguably stem from different analysis philosophies (see Hebart & Baker, 2018 for a detailed review)

Univariate analyses have dominated neuroimaging research for decades. Their focus is on studying brain activations from an activation-based philosophy. That is, the question that a researcher aims to respond with these analyses regards whether the brain is more or less active during certain conditions, assuming that higher activation implies a greater contribution of that region to a specific cognitive process. If a neuron shows a higher fire rate, that neuron is interpreted to be more engaged in the condition of study. Similarly, more BOLD signal on an fMRI study, larger voltage amplitudes on an ERP or more power in a certain frequency in a time-frequency analysis are all interpreted as involvement of a brain feature into the research condition (Hebart & Baker, 2018).

Conversely, multivariate analyses gained much popularity in the last two decades (Woolgar et al., 2016). They were popularized with the initial multi-voxel pattern analyses (Haxby et al., 2001), and later representational similarity analyses (Kriegeskorte et al., 2008) for fMRI data. Crucially, both of these methodologies can be applied to a diverse set of neuroimaging methodologies such as M/EEG (Grootswagers, Wardle, et al., 2017; King & Dehaene, 2014). Unlike univariate methods, multivariate analyses are usually framed within information-based philosophies, in which any difference in the patterns of activity between two variables (e.g. voxel similarity, variance or electrode voltage patterns) can be used to argue that two variables are different in a certain condition, and consequently, that this condition is being represented in the brain. This feature leads to the assumption that multivariate analyses are more sensitive to subtle information changes, since they combine data from multiple sources of variability (e.g. voxels or electrodes; Hebart & Baker, 2018). This has made them fundamental to study brain activity in instances where little activity changes are predicted across different conditions, such as anticipation of different objects (e.g. González-García et al., 2016; Peñalver et al., 2023; Stokes et al., 2009).

1.3. Paradigms and behavioral relevance of preparation

Among the overwhelming number of psychophysics paradigms than can be used in cognitive science only a few are suited to study anticipations. These should 1) elicit anticipatory activity, by giving information about incoming stimulation, 2) be flexible, so that they can accommodate a number of different tasks with minimal change to the design and 3) induce a behavioral response that allows testing the effectiveness of being prepared. Cueing (or cue-target) paradigms have been used in a large number of experiments with different purposes, which makes them ideal to study anticipatory processing (Battistoni et al., 2017; Feuerriegel et al., 2021). In these tasks, cues usually advance some characteristic of the target that helps with its detection or categorization. This characteristic can be the spatial location (e.g. Posner, 1980; Soto & Blanco, 2004; Theeuwes & Van Der Burg, 2007), orientation of a grating (Kok et al., 2012, 2017), or the category to which stimuli belong (Stein & Peelen, 2015), among many others. In this section we outline some of the behavioral effects that have been found in two contexts using this paradigm, depending on the type of information that cues provide: probability or relevance.

Overall, when the cue predicts with a certain probability any characteristic of incoming stimulation, the results consistently show faster reaction times and higher accuracies when said prediction is correct, or valid (Stein & Peelen, 2015; Posner, 1980; Kok et al., 2012; but not always, see Kok et al., 2017). This has been interpreted as increased perceptual sensitivity to cued targets (Rungratsameetaweemana & Serences, 2019).

In selective attention, cues give information about relevant target characteristics, such as location, orientation, color, or motion direction (Battistoni et al., 2017). Although there is a large body of research that attempts at measuring attention by manipulating the probability of certain target features, those results might be conflated with expectation effects (see below). Here, we only show studies where effects cannot be attributed to probability. In these studies, it is complicated to observe effects on behavioral performance measures, since uncued (ignored, unselected or unattended) trials usually receive different (or no) responses than cued ones. Indeed, several studies employ paradigms where cueing is used to elicit a neural attentional response, but behavioral performance acts as a correlate of attention without cueing instances to compare (Martinez-Trujillo & Treue, 2004; Treue & Maunsell, 1999). For instance, Stokes et al. (2009) designed a task where a cue gave information about a relevant target, either an "X" or an "O", and participants responded whether the following stimulus matched the information given by the cue. This was used to elicit neural anticipatory patterns for the attended stimulus, and behavioral accuracy was reported as evidence that participants were sufficiently engaged in the task. Similarly, Scolari et al. (2012) designed a task in which a central cue would indicate to respond to the orientation of a grating, its contrast or the identity of a letter. All trials where perceptually equated regardless of the attended feature, and behavior across the three conditions was only compared to ensure that all were equally difficult. Other studies that study attentional selection usually focus on what characteristics of the cue better help to select certain stimuli. Research suggests that when individuals are cued with prior information about a particular stimulus feature (vs. when the cue is not present, or it is uninformative), it improves their ability to discriminate and identify that feature effectively. For instance, (Yeshurun & Carrasco, 1999) manipulated the relative distance of a spatial cue with respect to the target where it was most effective. In their experiment the cue could appear in one out of four quadrants, and although the target would always appear in the selected quadrant, they manipulated the cues' excentricity, and found that the closer the cue was to the real location of the target, the better the response. Liu et al. (2007) showed that attention to locations is deployed faster than attention to certain features, and that both benefit from longer cue-target intervals. In another example, Wolfe et al. (2004) showed that in challenging visual search tasks where a target needs to be distinguished from similar distractors, prior knowledge of the target's features enhances performance if the cue is identical to the target vs. when the cue is a descriptive word.

Altogether, cue-target paradigms have proven to be effective to elicit anticipatory estates, although the behavioral consequences depend greatly on the type of information that cues provide.

1.4. How similar is preparation to perception?

As presented above, there is a large body of research that suggests that anticipatory and sustained representations take place in perceptual regions tuned to the prepared stimulus (Lee et al., 2013; Sreenivasan et al., 2014; Stokes et al., 2009; Giesbrecht et al., 2006; Peelen & Kastner, 2011; Soon et al., 2013; Esterman & Yantis, 2010; González-García et al., 2018). However, it is not clear whether these representations are similar to perception of the actual stimulus. Noteworthy, the study of the representational characteristics of top-down maintenance and perception has gone hand by hand with multivariate analyses. One of particular relevance is the canonical template tracking (CTT) procedure (Palenciano et al., 2023) that employs an independent task to assess through cross-classification the level of similarity between specific cognitive features (e.g. anticipation, motor preparation, imagery) and perception.

A field that has put significant efforts into solving this question is the study of visual imagery. For example, Cichy et al. (2012) demonstrated that imagery and perception share representations by training a classifier during an imagery task and then testing it during stimulus perception. Significant generalization was interpreted as evidence of similar representations. Relatedly, Dijkstra et al. (2017) found that the vividness of imagery depended on the overlap with perception in visual regions. Moreover, they later showed (Dijkstra et al., 2019) that although the regions involved in both domains are similar, the functional connectivity measured with dynamic causal modelling differed. Particularly, the inferior frontal cortex modulated the fusiform gyrus in both perception and imagery, but had a stronger modulatory effect in the latter. In addition to these results, other models propose different underlying mechanisms for each. Essentially, perception would act mostly through bottom-up cortical excitatory mechanisms, while mental imagery would use top-down inhibitory connections as a mechanism to construct perception-like visual representations (Koenig-Robert & Pearson, 2021; Pace et al., 2023).

Regardless of specific cognitive process, most models contend that top-down signals arise from associative regions (such as FPN in WM) and later induce specific representations in downstream sensory cortices. But how is bottom-up and top-down information represented in the same regions simultaneously and still maintain a similar representational structure? An accepted explanation is that information travels through different cortical layers (Self et al., 2019), which allows two signals with different origins to converge in the same brain region. These layers establish two types of connections. On the one hand, there are short connections between the layers within the same cortical column, which allow them to communicate with each other. On the other hand, they have long-range connections, at a greater distance, with neurons of the same layer in different regions of the brain. Specifically, bottom-up feedforward sensory information moves primarily through the central layers, while top-down feedback signals move through the peripheral ones. Studies employing laminar fMRI (Lawrence et al., 2019), have shown that perceived information can be decoded from central layers in V1 and V2, while non-perceived but contextual information (relevant, likely, or remembered) is decoded from external cortical layers (Muckli et al., 2015; Ortiz-Tudela et al., 2023; Petro & Muckli, 2017). Thus, top-down anticipatory feedback information could be represented in the same regions as bottom-up feedforward signals, but in different cortical layers.

So far, we have shown that top-down anticipatory representations may take place in similar regions as perceived information, they appear in several cognitive contexts and seem to be enacted through feedback connections from higher order regions. However, an essential question is whether anticipated information is represented in a similar manner regardless of the context, or whether contextual factors also modulate anticipatory preparation.

1.5. One or different forms of representing perceptual information? Attention vs. Expectation

Among the several top-down factors that influence perceptual processing, from anticipations to early sensory activity, two are of special relevance. On the one hand, some stimuli are more relevant to our behavioral goals than others. In the example we showed at the beginning of this introduction, we may need to find a taxi on a rainy day. On the other, we have prior knowledge about the probability of a certain stimulus appearing in a specific context. For instance, we expect to find a line of taxis near the airport entrance. The first context is related to selective attention, while the second one refers to perceptual expectations (Rungratsameetaweemana & Serences, 2019; Summerfield & Egner, 2016). Although both phenomena are closely related (Summerfield & Egner, 2009), the study of attention has been prominent in cognitive science since the early beginnings of the field, such as the well-known quote by William James “Everyone knows what attention is” (James, 1890). On the contrary, although the interest on expectations also started at the beginnings of experimental psychology, with von Helmholtz in 1867 coining the term “unconscious inferences” to describe the effect of learned regularities on visual perception, this notion lost momentum, and for decades the terms attention and expectation ended up being used more or less interchangeably (Summerfield & Egner, 2009). For instance, the classical Posner attention cueing paradigm (Posner, 1980) measures attention by manipulating the expectations of where a target might appear. However, the study of expectations regained interest in the last years (Friston, 2005), and recent accounts dissociate the two terms. Attention refers to features that are behaviorally relevant, without being more likely, and expectation refers to features that are likely to occur, regardless their particular behavioral relevance (Summerfield & Egner, 2009). Understanding their influences in anticipatory representations is crucial to understand whether preparation is a unified or heterogeneous mechanism. Hence, in this last section we introduce the two terms, their effects during anticipations and target processing, the main theoretical models that have attempted to describe them, and previous research that has studied their interactions and differences.

1.5.1. Selective attention

The effects of preparatory attention have been studied in a variety of modalities, from single cell recordings in monkeys, to M/EEG and fMRI in humans. For example, in two single cell studies, Chelazzi et al. (1993, 1998) recorded inferotemporal (IT) neurons while monkeys performed memory-guided visual search tasks with complex stimuli. Neurons exhibited higher activity during the delay period following cues indicating behaviorally relevant objects that they were tuned to, reflecting the maintenance of specific contents in preparation for visual search. Human M/EEG studies have illustrated the role of alpha activity in anticipatory attention. Snyder & Foxe, (2010) showed alpha increases in color and motion selective cortices when the specific feature was unattended. A similar result was found for faces and oriented gratings (Jokisch & Jensen, 2007). These results have been interpreted as alpha reflecting the top-down suppression of irrelevant or distracting information (Battistoni et al., 2017). Research using fMRI expanded on these findings, revealing feature-specific increases in neural activity during preparatory phases. For instance, cues directing attention to color or motion led to corresponding activity increases in the relevant brain regions (Chawla et al., 1999). Stokes et al. (2009) found that it was possible to classify to what letter ("X" vs. "O") participants were cued to attend. In addition, (Peelen & Kastner, 2011) found a similar effect in naturalistic scenes, and crucially, were also able to correlate the results of prestimulus classification in object selective cortex to subsequent behavioral performance.

During target processing, evidence often suggests that attended stimuli elicit increased activations (Rungratsameetaweemana et al., 2018). EEG research has demonstrated larger voltage amplitudes for attended information compared to ignored one in parietal and frontal electrodes in early latencies (around 200 ms), such as positions (He et al., 2004) or words (Perrone-Bertolotti et al., 2020). Additionally, various oscillatory patterns have been linked to attention: increased Theta (4-8 Hz) power in frontal and medial electrodes in situations requiring cognitive control, such as distractor inhibition (Cavanagh & Frank, 2014), decreased Alpha (8-12 Hz) power and synchrony for selected stimuli in occipital and parietal electrodes (e.g. Bagherzadeh et al., 2020; Banerjee et al., 2011; Feng et al., 2017), and enhanced Gamma power (Gruber et al., 1999) and synchrony (Doesburg et al., 2008) also for selected stimuli in occipital electrodes between 200 and 300 ms. Studies employing fMRI have frequently observed increased cortical activity for relevant stimuli in occipitotemporal and parietal regions in both non-human (Bichot et al., 2005; Burrows et al., 2014; Chelazzi et al., 1998; Cohen & Maunsell, 2011) and human (Kastner et al., 1999; Peelen & Kastner, 2011; Serences et al., 2004) primates.

Altogether, these findings support frameworks such as the biased competition theory (Beck & Kastner, 2005; Desimone & Duncan, 1995). This model argues that, when there are different stimuli in the sensory field, they compete for representational resources. Top-down selection increases cortical sensitivity to the most relevant stimuli, while suppressing activity of neurons tuned to the ignored ones. These biases happen through cortical templates, which are "short-term descriptions" (Desimone & Duncan, 1995, pp. 199-200) of the relevant features. These templates would be generated in prefrontal regions during the prestimulus period, are held in WM, and induce top-down biases in downstream regions, thus increasing sensitivity to relevant stimuli.

1.5.2. Expectation

Cues providing probabilistic information also lead to the activation of specific templates that later bias further processing. Kok et al. (2014) found stimulus-specific activity in V1 when a stimulus was expected, but not shown. Moreover, they later used EEG (Kok et al., 2017) and found that it was possible to decode the most probable orientation of a gabor after it had been predicted by an auditory cue. Similar results have been found for direction (Ekman et al., 2017), motor patterns (de Lange et al., 2013) or abstract shapes (Hindy et al., 2016).

However, contrary to attention, it is a common finding that during target processing perception of probable stimuli induces activity decreases (or, conversely, unexpected information induces increases), a phenomenon often labeled "expectation suppression". EEG correlates evoked by probable information, such as the mismatch negativity (Garrido et al., 2009; Pazo-Alvarez et al., 2003), are thought to reflect its reduced processing compared to unexpected one. Similar effects where novel stimuli generate larger EEG amplitudes than expected ones have also been found (e.g. Manahova et al., 2018). Likewise, oscillatory activity appears to reflect coincident effects. Such is the case for midfrontal Theta power, associated with inhibition and control (Cavanagh & Frank, 2014), which increases in contexts of unexpectedness in oddball and reinforcement learning paradigms (Cavanagh et al., 2012). In addition, Gamma power has been found to decrease when similar stimuli are repeated over time (Brunet et al., 2014). Relatedly, studies using fMRI have found reduced activations for probable target stimuli in regions such as V1 and lateral occipital cortex (e.g. Kok et al., 2012; Walsh & McGovern, 2018; see Feuerriegel et al., 2021 for a detailed review).

One of the strongest theoretical frameworks that attempts at explaining how prior probabilistic knowledge influences cognitive processing is the predictive processing framework (Friston, 2005). This model stems from the aforementioned idea of cortical organization, and has a great influence on the study of preparation in general and expectations in particular. The model proposes that at every moment the brain acts as a Bayesian predictor. Previous knowledge about statistical regularities would serve as prior, which travels top-down through peripheral cortical layers to downstream areas, facilitating perception of probable stimuli. Simultaneously, bottom-up signals would travel from sensory organs through central layers, where inter-layer signals will be influenced by and influence predictive information. The mismatch between predicted signals and bottom-up information would generate "prediction errors", which are theorized to update predictions, generating posteriors that can later on act as priors and flexibly adapt top-down learning and facilitate responses (de Lange et al., 2018).

This fits an explanation of the effects of anticipations and expectation suppression. Top-down signals would preactivate neurons tuned to probable stimuli in sensory regions in the prestimulus interval. Then, if the prior is correct, bottom-up signals would carry small prediction errors, which would induce little activation. Conversely, false expectations would need more activity to update top-down knowledge, and thus lead to increased activations. However, the nature of the precise neural mechanisms that generate expectation suppression is still debated. There are two main and, in principle, mutually exclusive explanations (de Lange et al., 2014; Press et al., 2020). The first one proposes that the effect is a consequence of the suppression of neural responses that are similar to top-down predictions, and therefore, redundant (Murray et al., 2004; Friston, 2005). This results in a "dampening" of neural populations tuned to predicted stimuli, and hence poorer representations. The second explanation proposes that the decrease in activation is explained by a decrease of activity of neurons tuned away from expected information, and a “sharpening” of the representation, thus leading to better representational activity (Kok, Jehee, et al., 2012). Both explanations have received support from the literature (González-García & He, 2021; Kok, Jehee, et al., 2012; Press et al., 2020; Richter et al., 2018; Walsh & McGovern, 2018). Crucially, this latter mechanism has been related to findings in selective attention that point to sharper representations for selected/relevant stimuli in frontoparietal and perceptual regions (Goddard et al., 2022; Woolgar et al., 2015). Hence, category-selective neurons would increase their sensitivity to both relevant or probable categories.

1.5.3. Interactions and differences between relevance and probability

Given their close relation, a growing body of literature has studied how attention and expectations interact, and, crucially, how they differ. In this last section we briefly review the studies that have attempted at examining the effects that attention and expectation have on each other, and finally, what evidence accounts for the notion that they are indeed different orthogonal processes.

The studies that have been carried out to examine the interaction between relevance and probability often lay under the predictive processing framework. It considers two possible theories that account for the effects of attention. The first one proposes that attention acts as a filter, sharpening the activity of populations tuned to probable features or locations, increasing neural gain in such populations. Alternatively, attention might increase the strength of predictions, increasing the weight of sensory evidence after target perception when probable stimuli match relevant features (Feldman & Friston, 2010; Summerfield & Egner, 2014). Evidence has mostly supported this last view (Kok, Rahnev, Jehee, Lau, & De Lange, 2012; Summerfield & Egner, 2016). For instance, Kok, Rahnev, et al. (2012) factorially manipulated relevance and probability, by having blocks in which information was more likely to appear in a specific location, and attentional cues in each trials indicating the relevant location to respond. They found that attended probable trials reversed the expectation suppression effects, and indeed where the ones where activations where largest. In the same line Jiang et al. (2013) sought to examine whether attention sharpened sensory representations and how this interacted with expectations. They found the classification between expected and unexpected stimuli was largest when they where attended. This has been demonstrated several times (Alilović et al., 2019; Auksztulewicz et al., 2017) although there is also conflicting evidence (Ekman et al., 2017; Yon et al., 2018). An example of competing evidence comes from Richter & de Lange (2019). Using a statistical learning paradigm, they showed that expectation suppression effects were increased for attended stimuli. They justified this difference with previous literature by stating that attention might influence expectations differently when there is a competing stimulus or task that can effectively draw attention away, compared to stimuli that although task-irrelevant, they might capture some attention and bias possible effects. Altogether, although it is clear that relevance and probability interact during perceptual processing, the exact relationship is still debated.

Finally, some studies have sought to directly contrast relevance and probability effects (Simon et al., 2018). Wyart et al. (2012) orthogonally manipulated attention and expectation in a behavioral task in which two colored placeholders located in their left and right visual fields could contain or not a grating. A retro-cue indicated the relevant placeholder to respond to. Before stimulus presentation, a relevance cue indicated the location that was most likely to be probed, while a probability cue indicated in which placeholder the grating was more likely to appear. They found that although both relevance and probability increased detection sensitivity, attention cues where most effective in cases of signal present trials, while probability cues were effective for signal absent trials. This was interpreted as attention functioning as a noise reduction mechanism by enhancing error signals, while expectation could boost inputs by sharpening probable stimuli. In another example Zuanazzi & Noppeney (2019) found activation in the FPN for selected visual and auditory targets, but no effect for expected visual targets. Meanwhile, Gordon et al. (2019) linked selective attention to lower levels of the visual hierarchy, and expectations to more complex instances of visual perception. Crucially, all of these studies have focused on the effects that attention and expectation have during target processing, but have not contrasted how they influence anticipatory representations.

Chapter II

GOALS AND HYPOTHESES

In the introductory section of this thesis, we have outlined the nature of endogenous brain processing and, in particular, the role of anticipatory representations in top-down functioning. Even though foreknowledge has been frequently studied, the extent to which the anticipation of specific information conforms a whole uniform mechanism or a heterogeneous set of neural functions that depend on the informational role of the content anticipated and the cognitive context is currently unclear. Hence, understanding the spatiotemporal dynamics of preparation across contexts is crucial to further understand proactive cognition, which is a key aspect of human behavior. Moreover, we have highlighted two well-known processes that are known to include anticipatory components: selective attention and perceptual expectations. Although there is an ongoing debate not only about how they interact, but also about whether they are actually indistinguishable from one another, literature has seldom directly compared them. Moreover, in the cases where these have been contrasted, it has been at target perception and response stages. Thus, knowledge of the commonalities and dissimilarities of their preparatory stages, and their link with effects observed during target processing, is missing.

The **overall goal** of this thesis is to advance the knowledge about the nature of neural top-down preparation by studying whether it is a uniform or a context-dependent brain function. We first addressed this with two parallel neuroimaging studies in which we studied and contrasted preparation in relevance and probability contexts, using EEG and fMRI. Then, in a third study we analyzed target processing in the same paradigm to investigate how proactive mechanisms bias neural populations to efficiently process information. Specifically, the three studies aimed to answer the following specific **global questions**:

1) Study I (Chapter 3). What are the temporal dynamics of anticipation? Do preparatory representations of stimulus category adapt to contexts of anticipated relevance vs. probability?

2) Study II (Chapter 4). What are the key brain structures involved in category-specific anticipation? Is this preparatory information represented in the same regions as actual target perception? How do these change across relevance and probability contexts? Are anticipatory neural markers linked to behavioral efficiency?
3) Study III (Chapter 5). How does top-down anticipation influence neural populations tuned to specific categories during stimulus processing? Is an unexpected stimulus represented similarly to a goal-(ir)relevant one?

2.1. Differences between anticipatory representations across selective attention and perceptual expectations. Studies I and II.

The prerequisite goal **(G0)** was to design a task to study category-specific preparation by directly contrasting attention and expectation. This was composed of the following sub-goals. **G0.1.** To develop a cue-target experimental paradigm optimized to orthogonally manipulate selective attention (relevance) and expectation (probability). **G0.2.** To equate the two conditions at the perceptual and response levels. **G0.3.** To design the paradigm so that it allows performing univariate and multivariate analyses of the representational structure of anticipated contents for top-down preparation. **G0.4.** To adapt this paradigm to EEG and fMRI to maximize the particular strengths of each neuroimaging method and to perform joint analyses with the results of the two experiments.

2.1.1. Temporal dynamics of anticipatory representations. Study I

Our primary objective **(G1)** was to examine the timeline of the mechanisms involved in category anticipation and directly contrast them across contexts of relevance and probability. More specifically, we had the following sub-goals: **G1.1.** To analyze the influence of endogenous signals on the behavioral performance in two relevant cognitive contexts: selective attention and perceptual expectation. **G1.2.** To examine the temporal development of electrophysiological activity associated with the representation of stimulus categories in attention and expectation, by means of symbolic signals containing information about the stimulus category to be selected, or the most probable one, respectively. **G1.3.** To contrast and detail how evidence of information representation differs during anticipation in the two contexts, by applying a cross-classification approach to directly compare the neural activity patterns underlying the representation of relevant versus probable stimuli. **G1.4.** To examine the reinstatement of perceptual patterns during anticipation, by collecting data from an independent localizer and applying a Canonical Template Tracking approach.

Following these goals and the significant differences in target processing observed in previous studies (Jiang et al., 2013; Kok et al., 2012; Summerfield & Egner, 2009; Wyart et al., 2012), our main hypothesis **(H1)** was that the neural coding responsible for these divergent outcomes should be distinguishable at the preparatory stage. Specifically, we had the following hypotheses: **H1.1.** Probable stimuli would elicit faster and more accurate responses than non-probable ones. We did not expect differences between cued and uncued trials in attention. **H1.2.** Both manipulations would lead to the preactivation of anticipated contents, evidenced by significant anticipatory decoding of the two target categories. **H1.3.** Attention and expectation would not cross-classify from one condition to the other, and if they did so, it would be at different latencies measured by a temporal generalization analysis. **H1.4.** The level of perceptual reinstatement would differ between the two conditions.

2.1.2. Brain regions involved in specific anticipatory representations across domains. Study II

Our main goal **(G2)** was to study neural populations involved in anticipatory representations of stimuli in attention and expectation contexts, with the following subgoals: **G2.1.** To replicate the behavioral results of Experiment I. **G2.2.** To explore the regions that are differentially recruited during anticipation in relevance and probability contexts. **G2.3.** To analyze the similarity between anticipated information in attention and expectation, by performing a cross-classification analysis. **G2.4.** To examine the level of perceptual reinstatement between anticipation and target perception. **G2.5.** To study the behavioral relevance of anticipatory representations.

Our overarching hypothesis **(H2)** was that both manipulations would lead to context-specific preactivations. We had the following specific hypotheses: **H2.1.** Behavioral performance would be similar to Experiment I. **H2.2.** Specific anticipatory states would be decoded in various brain regions in both contexts, including in category-selective regions of the ventral stream. **H2.3.** Classification would not generalize across attention and expectation, indicating partially distinct neural mechanisms. **H2.4.** Preactivations would exhibit similarities with target decoding, with these being more pronounced in the attention condition. **H2.5.** Decoding would be positively correlated with behavioral performance.

2.2. Top-down modulations of target stimulus processing in attention and expectation. Study III

In this final study, we sought **(G3)** to explore the influence of top-down mechanisms on the neural activity elicited by the perception of relevant and probable stimuli, and to investigate whether such anticipation is consistent with either dampening or sharpening accounts. To that aim, we re-analyzed the data from the two previous experiments focusing on target locked activity, following parallel analysis rationales for each of them. The specific sub-goals were the following: **G3.1.** To find evidence of expectation suppression for predicted targets, and attention enhancement for relevant stimuli, by contrasting univariate activations across cued and uncued trials in attention and expectation. **G3.2.** To adjudicate between dampened vs. sharpened neural responses in attention and expectation, by comparing the classification accuracies to cued and uncued targets in both conditions. **G3.3.** To study the temporal stability of the effects and to examine whether generalization to preparation interval varies across cueing conditions. **G3.4.** To investigate the involvement of frontoparietal regions in potential dampening and sharpening effects. **G3.5.** To explore whether attention and expectation affect target perception differently or, conversely, whether attention is automatically directed to either probable or surprising stimuli.

Our general hypothesis **(H3)** was that dampening and/or sharpening effects would be different across selection and probability conditions, with the following specific hypotheses: **H3.1.** Selected targets would elicit larger activations, while probable stimuli would show reduced activations compared to improbable ones. **H3.2.** Attended targets would elicit sharper representations, evidenced by increased decoding for relevant vs. irrelevant stimuli. **H3.3.** In the EEG experiment, sharpened representations would generalize worse across different trials and to the anticipatory window. Conversely, dampened representations should show better generalization to the cue and different trials. **H3.4.** Visual areas and frontoparietal regions would exhibit evidence of dampening and sharpening effects. **H3.5.** Altogether, attention and expectation would elicit a different pattern of dampening and sharpening associated to each condition.

Chapter III
Study I

The content of this chapter is published as Peñalver, J. M. G., López-García, D., González-García, C., Aguado-López, B., Górriz, J. M., & Ruz, M. (2023). Top-down specific preparatory activations for selective attention and perceptual expectations. *NeuroImage*, *271*, 119960.
https://doi.org/10.1016/j.neuroimage.2023.119960

3.1. Summary

Proactive cognition brain models are mainstream nowadays. Within these, preparation is understood as an endogenous, top-down function that takes place prior to the actual perception of a stimulus and improves subsequent behavior. Neuroimaging has shown the existence of such preparatory activity separately in different cognitive domains, however no research to date has sought to uncover their potential similarities and differences. Two of these, often confounded in the literature, are selective attention (information relevance) and perceptual expectation (information probability). We used EEG to characterize the mechanisms that pre-activate specific contents in attention and expectation. In different blocks, participants were cued to the *relevance* or to the *probability* of target categories, faces vs. names, in a gender discrimination task. Multivariate Pattern (MVPA) and Representational Similarity Analyses (RSA) during the preparation window showed that both manipulations led to a significant, ramping-up prediction of the relevant or expected target category. However, classifiers trained on data from one condition did not generalize to the other, indicating the existence of unique anticipatory neural patterns. In addition, a Canonical Template Tracking procedure showed that there was stronger anticipatory perceptual reinstatement for relevance than for expectation blocks. Overall, results indicate that preparation during attention and expectation acts through distinguishable neural mechanisms. These findings have important implications for current models of brain functioning, as they are a first step towards characterizing and dissociating the neural mechanisms involved in top-down anticipatory processing.

3.2. Introduction

For decades, research in cognitive psychology has studied behavior while manipulating external factors, which led to theoretical models that framed cognition mostly from a reactive point of view. Recent years have witnessed a renaissance of proactive cognition, where endogenous top-down mechanisms play a core role in brain functioning. Within this framework, preparation is conceptualized as an endogenous function that takes place in anticipation of incoming inputs or demands and improves subsequent behavior (Battistoni et al., 2017; González-García et al., 2016). Neural preparatory activity has been shown for a plethora of processes, including attention (Battistoni et al., 2017; Kastner et al., 1999; Nobre & Serences, 2018), expectation (Aranda et al., 2010; Kok et al., 2017), working memory (Koshino et al., 2015; van Driel et al., 2017), or cognitive control (Baines et al., 2011; González-García et al., 2016; Hebart & Baker, 2018). Similarly, influential models have proposed different ways in which anticipatory patterns interact with stimulus inputs to guide perception. Examples of this are the Predictive Coding (Auksztulewicz & Friston, 2016) or the Biased Competition frameworks (Desimone & Duncan, 1995). However, currently these models are silent as to how different preparatory phenomena relate to each other, and whether they reflect common or diverging underlying top-down mechanisms.

Selective attention and perceptual expectation are complex functions that involve top-down and bottom-up elements. Attention refers to the selection of *relevant* information based on specific goals (Nobre & Serences, 2018), while expectation involves predictions based on prior *probability* (Schröger et al., 2015). Studies of selective attention have manipulated the relevance of information using cues that indicate the stimulus or dimensions to respond while ignoring others (Battistoni et al., 2017; Hong et al., 2017; Nobre & Serences, 2018; Stokes et al., 2009). Expectation has primarily (but not only, see e.g. Summerfield & De Lange, 2014) been manipulated with cues that inform about the most probable stimulus (de Lange et al., 2013; Kok et al., 2017; Wyart et al., 2011). Previous research focusing on the effect of attention and expectations on target processing has shown diverging (brain activity) and overlapping (behavioral) results. Although both lead to behavioral improvements (Ho et al., 2012; Stein & Peelen, 2015), several studies reveal that they can be, at least partially differentiable (Rungratsameetaweemana & Serences, 2019; Summerfield & Egner, 2009). Neuroimaging studies so far have found differences in contexts of relevance and probability (see Summerfield & Egner, 2016), including activity increases for selected target stimuli and decreases for expected ones. Another fruitful field of research has focused on how they interact, showing in some cases how selective attention can modulate the effect of expectations (Alilović et al., 2019; Auksztulewicz et al., 2017; Jiang et al., 2013; Kok et al. 2012) and a lack of interactions in others (e.g. Ekman et al., 2017; Yon et al., 2018).

Studies that have tried to overcome frequent confounds between attention and expectation (e.g. Posner, 1980; Schröger, 1996) have shown separate roles of relevance and probability during target processing (Auksztulewicz et al., 2017; Gordon et al., 2019; Simon et al., 2019; Wyart et al., 2011; Zuanazzi & Noppeney, 2019). On the other hand, research focused on preparatory activity of either attention or expectation has provided seemingly overlapping results. Cues indicating relevance in selective attention (Battistoni et al., 2017; Nobre & Serences, 2018) preactivate relevant regions of space processing (Giesbrecht et al., 2006), specific shape patterns in visual cortex (Stokes et al., 2009), patterns in category (Esterman & Yantis, 2010; González-García et al., 2018) and object-selective perceptual regions (Peelen & Kastner, 2011; Soon et al., 2013). Similarly, cues providing probabilistic information lead to the preactivation of specific templates of oriented gabors (Kok et al., 2017), direction (Ekman et al., 2017), motor patterns (de Lange et al., 2013) or abstract shapes (Hindy et al., 2016). However, all these previous investigations are agnostic regarding the potential similarities or differences in such top-down preparation across relevance and probability anticipation. Unraveling the differences in how anticipatory activity in different contexts reflects the upcoming information is a necessary step to understand the differences between attention and expectation, and is also essential for theoretical models that explain the neural basis of these two phenomena (Auksztulewicz et al., 2018; de Lange et al., 2018; Desimone & Duncan, 1995).

In our study, we employed different multivariate pattern analyses of EEG data to directly compare the representational structure of anticipated contents for top-down preparation in contexts capitalizing on selective attention (relevance) or expectation (probability). That is, we examined whether the coding of anticipated stimulus content differs on the basis of the type of anticipation. To do so, we embedded a sex/gender discrimination in a cue-target paradigm. Here, depending on the block, cues provided information about the upcoming relevance or probability (Egner et al., 2010; Wyart et al., 2011) of face or name stimulus categories. In addition, we ran an independent localizer to study similarities between preparation and perception across contexts. Our first goal was to study the anticipatory mechanisms during preparation by means of time-resolved representational similarity analysis (RSA, Kriegeskorte, 2008). Then, we used multivariate pattern analysis (MVPA, Grootswagers, Wardel, et al., 2017) to examine whether anticipated stimulus categories are represented with differential fidelity during selected compared to probable targets. Next, we used a cross-classification approach to directly contrast the patterns of activity underlying the representation of relevant vs. probable stimuli. Taking into consideration the striking differences observed during target processing in these contexts (Jiang et al., 2013; Kok, Rahnev, Jehee, Lau, & De Lange, 2012; Summerfield & Egner, 2009; Wyart et al., 2012), our hypothesis was that the neural coding that leads to such different consequences should be dissociable from an early processing stage. To further understand the proposed differences, we leveraged a Canonical Template Tracking approach (González-García et al., 2021; Palenciano et al., 2022; Wimber et al., 2015) to observe the extent to which preparation induces the reinstatement of overall perceptual information in each condition. Given the dissociations between attention and expectation observed during target processing (Gordon et al., 2019; Wyart et al., 2011) and the apparent commonalities reported during their top-down preparatory states (e.g. Battistoni et al., 2017), our overall hypothesis was that both manipulations would lead to the preactivation of anticipated contents but through at least partially different neural mechanisms.

3.3. Methods

Methods are reported in accordance with the Committee on Best Practices in Data Analysis and Sharing (COBIDAS) for M/EEG (Pernet et al., 2018).

3.3.1. Data and code availability

Original code has been deposited at Github and is publicly available as of the date of submission. Results have been deposited at OSF website. Raw data are available online at OpenNeuro.

3.3.2. Participants

Forty-eight participants (mean age = 22.06, range = 18-31; 29 women, 18 men, 1 non-binary) from the University of Granada were recruited and received from 20 to 25 euros, depending on their performance. Two additional participants were discarded due to low behavioral accuracy (less than 80%) or excessive noise in the EEG (more than 20% discarded trials). They were all native Spanish speakers, right-handed with normal or corrected vision, and signed informed consent prior to participation. Besides, to comply with COVID-19 guidelines, the temperature of participants was measured upon arrival (always <37ºC), they confirmed to have had no illness symptoms in the days prior to the experiment and wore a face mask during the whole session. Sample size was calculated to achieve a statistical power of 80% for an estimated small effect size (Cohen's d = 0.3) and three independent variables (block x category x cueing). Using PANGEA (Power ANalysis for GEneral ANOVA designs, Westfall, 2016) we obtained a minimum of 32 participants to be able to detect the block x cueing interaction in reaction times and behavioral accuracy, our main behavioral prediction. To fit the counterbalancing scheme, we tested 48 participants. This sample size provides an estimated power of 94% under the described parameters.

3.3.3. Apparatus, stimuli, and procedure

Stimulus presentation and behavioral data collection were done with The Psychophysics Toolbox 3 (Brainard, 1997) on MATLAB (v.2020) in a Microsoft PC. Stimuli were presented on an LCD screen (Benq, 1920x1080 resolution, 60 Hz refresh rate) over a grey background. We employed 160 male and female faces (50% each, with ~6ºx9º visual angle, extracted from The Chicago Face Database (Ma et al., 2015) plus 160 unique Spanish male and female names (50% each, with ~8ºx2º visual angle). Four different geometrical shapes (circle, square, rain-drop and diamond with thin black outlines, unfilled, ~2ºx2º visual angle) were used as cues in the main task. The sound stimuli employed in the localizer blocks consisted of four different tones (250, 300, 350 and 400 Hz).

The main task was a cue-target paradigm where, depending on the block, cues carried information about either the *relevance* (attention) or the *probability* (expectation) of upcoming face or word targets. Each trial started with the presentation of this visual cue. For each participant, and to avoid perceptual confounds, two cue shapes (counterbalanced across participants) were associated with faces and two with names. Importantly, cue pairs (the cue associated with faces and names) changed through the experiment. This way, the first cue for faces (e.g. a circle) appeared in half of the blocks with the first cue for names (e.g. a square) and the other half with the second cue for names (e.g. diamond). The task was to indicate the sex/gender of this target (male or female). Participants pressed one of two keys (“a”, “l”, counterbalanced across participants) to respond whether or not the target belonged to the gender stated at the beginning of each block. Half of the blocks belonged to the attention condition, and the other half to the expectation condition. Participants were verbally instructed to use the cues in the two blocks to respond as fast as possible while avoiding mistakes. At the beginning of each block (Figure 3.1B), they were informed about the block (attention or expectation), the target sex/gender ("Is the target male/female?"), and the two cues (one for faces and one for names). Importantly, and since attention and expectation are involved in almost any act of visual perception, we aimed at manipulating one process while keeping the other constant. In attention blocks, the cue indicated the *relevant* stimulus category to select (either faces or names). Only if the stimulus belonged to the relevant category (50% trials, cued), the participant had to perform the gender discrimination task on the target. Otherwise, participants had to answer '*no*' regardless of the stimulus sex/gender (non-relevant category, uncued). Note that this manipulation of relevance, where further processing has to be applied only to selected stimuli, is similar to that employed in previous literature (e.g. Baldauf & Desimone, 2014; Saenz et al., 2002; Summerfield et al., 2006; Womelsdorf et al., 2006). Therefore, participants had to give an answer to all attention trials, and had to be prepared to perform the gender judgment task. Importantly, both relevant and non-relevant targets were matched in expectation, as by design they appeared with a 50% probability after each attention cue. On the other hand, in expectation blocks the cue indicated the *probable* category of the target, with a 75% likelihood (e.g. de Lange et al., 2013; Kok et al., 2017) for similar manipulations). Here, participants had to perform the gender discrimination task in all trials, whether or not the target was cued. This way, both the expected and unexpected targets were equally relevant.

In every trial of the main task, the sequence of events was as follows: a 50 ms cue was followed by a fixed Cue-Target Interval (CTI) of 1500 ms and then the target appeared for 100 ms. Trials were separated by 2500 ms intervals. Auditory (tone, 400 Hz, lasting for 300 ms) and simultaneous visual feedback (words "attention" or "expectation", depending on the block presented for 500 ms) appeared in case of a wrong answer 1.3 seconds after target presentation, without altering the trials duration. Each trial lasted 4.15 seconds and each block 1.23 minutes. This main task was composed by 32 blocks of 20 trials each, or 640 trials in total.

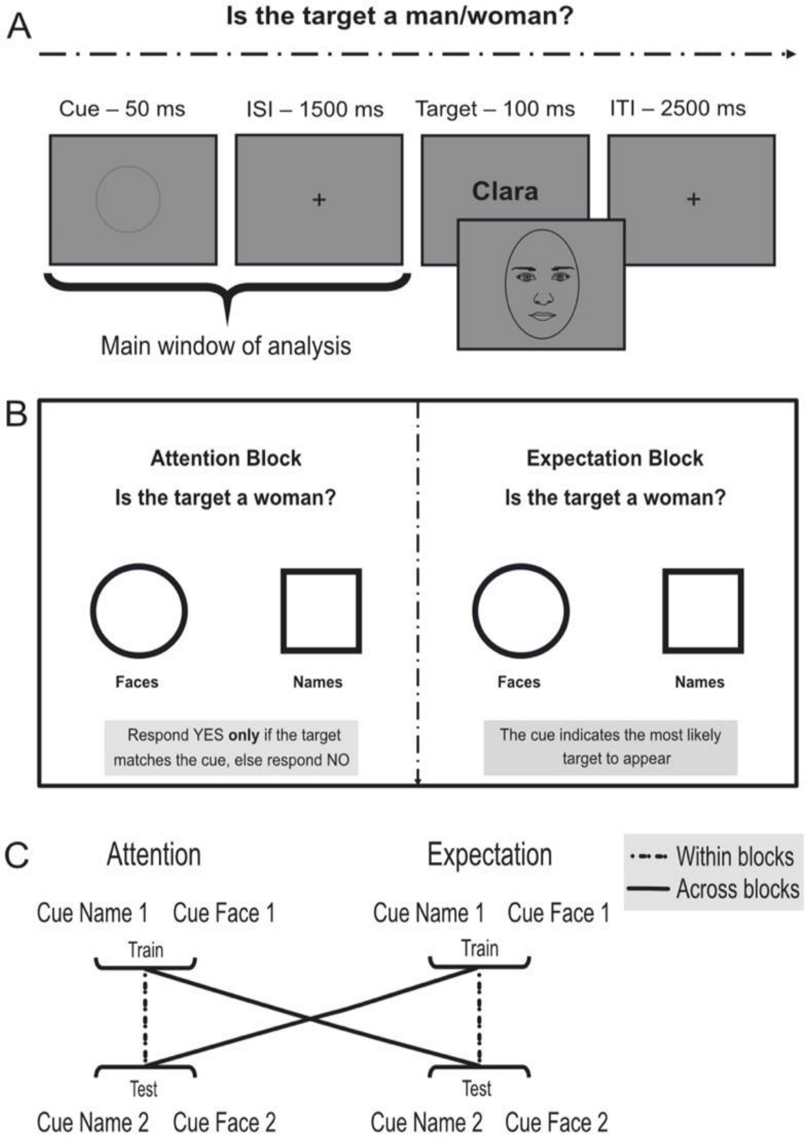


Figure 3.1. Behavioral task and design (A) Behavioral task: example trial. Participants were cued about an incoming target stimulus (a face or a name) with which they performed a sex/gender classification task. (B) Block starting screen. In attention blocks, the cue indicated the relevant stimulus. Participants performed the judgment on top of the screen only if the stimulus matched the cue. In non-relevant trials, participants responded with the "no" key. In attention blocks the cues were not predictive of the probability of the target (50% faces vs. names) or the response. In expectation blocks the cue indicated the probability of the stimulus category, which appeared 75% of times. Participants had to perform the judgment regardless of the cue (thus predicted and unpredicted targets were equally task-relevant), which carried no response information either. That is, in both blocks participants responded in all trials and perceptual details were fully equated. (C) MVPA classification rationale. In cue-locked within-block analyses, the classifier was trained to differentiate between anticipated (relevant or probable, depending on the block) faces vs. names with one pair of cues and then tested on the other, in either attention or expectation conditions. Note that classifiers could not use block information to differentiate between faces and words, as these were matched in every contrast. Across-blocks, the classifier was trained with one pair of cues within one condition (e.g., attention) and then tested on the other pair of cues within the other condition (e.g., expectation).

In addition, attention and expectation blocks were interspersed with localizer ones, used to measure target perceptual processing without overt motor activity (adapted from Egner et al., 2010). In these, the same stimuli as in the main task were presented. They were preceded by auditory cues that predicted either faces or names with 75% validity. Similar to the main task, there were two tones predicting faces, and two tones predicting names. One of each type was used per block. This manipulation was not used for the analyses reported in this work. Participants had to press a key *only* if the stimulus appeared upside down (rotated 180º, 10% of trials equated in stimulus category and sex/gender), regardless of cue validity. At the beginning of each block, a screen indicated the type of block and the two tones for the block. There were 16 localizer blocks of 40 trials each, 640 trials in total. The cue lasted 200 ms, the visual stimuli appeared after an ISI of 1500 ms and stayed onscreen for 100 ms. Trials were separated by 1500 ms intervals. Overall, each localizer trial lasted 3.3 seconds and each localizer block lasted 2.12 minutes.

The three types of blocks (attention, expectation and localizer) appeared in a fully counterbalanced order, so that that they preceded and followed each other an equal number of times. Cues and target stimuli were also fully counterbalanced across participants. In total, the whole experimental session lasted approximately 80 minutes, with additional practice and EEG preparation time.

3.3.4. EEG data acquisition and preprocessing

Acquisition

High-density EEG was recorded with 64 channels mounted on an elastic cap (actiCap Slim, BrainVision) at the Mind, Brain and Behavior Research Center (CIMCYC) of the University of Granada. Impedances were kept below 10 kΩ as recommended by the amplifiers' manufacturers. EEG activity was referenced online to FCz and recorded at a sampling rate of 1000 Hz.

Preprocessing

Data were preprocessed using EEGLAB (Delorme & Makeig, 2004) and custom MATLAB scripts (López-García et al., 2022). EEG recordings downsampled to 256Hz, digitally low-pass filtered using FIR filter at 120 Hz and high-pass filtered at 0.1 Hz. A notch bandpass was applied at 50 Hz and 100 Hz to remove line noise and its harmonics. Channels with excessive noise were identified by visual inspection and removed from the data (1.85% of channels on average, range 0-5). We used different epochs for cue and target stimuli and in each we split the data into 3 seconds epochs (-1 to 2 seconds after the onset of each stimulus). Then, Independent Component Analysis (ICA) was carried out to remove eye artifacts (i.e. blinks and lateral eye movements) using the *runica* algorithm from EEGLAB. Component rejection was guided by visual inspection of scalp maps, raw activity and power spectrum. *ICLabel* was used for further confirmation. In average 1.85 components were removed per participant (range 2-4). Next, we performed automatic trial rejection to prune the data from non-stereotypical artifacts. It was based on three factors: 1) abnormal spectra: trials in which the spectrum deviated from baseline by ±50 dB in the 0–2 Hz frequency window (sensitive to remaining eye artifacts) or deviated by −100 dB or +25 dB in 20–40Hz (sensitive to muscle activity); 2) improbable data: the probability of occurrence of each trial was computed by determining the probability distribution of values across trials, with a rejection threshold established at ±6 SD; (3) extreme values: all trials with amplitudes in any electrode out of a ±150μV range were automatically rejected (see Keil et al., 2014; López-García et al., 2020, 2022 for similar preprocessing routines). The three methods in sum yielded an average of 8% of rejected trials per participant (range 1.8%-19%). Afterwards, the removed channels were repaired by spherical interpolation. We then applied common average to re-reference the data, given its widespread use and its optimal adaptation to high-density recordings (Pernet et al., 2018). Finally, trials were baseline corrected in the -200 to 0 ms prior to stimulus onset.

3.3.5. Analyses

Behavioral

The main task design had three within-subject factors: Block type (attention vs. expectation), Cueing (cued vs. uncued) and Stimulus category (faces vs. names). We calculated three-way repeated measures ANOVA separately for behavioral accuracy and reaction times (RTs) employing JASP (Love et al., 2019). For each participant, trials with longer or shorter RTs than the average ± 2 SDs were discarded (11.54% on average). Behavioral results in the localizer were not relevant and were just considered as exclusion criteria in case of poor performance.

Representational Similarity Analysis

Representational Similarity Analysis (RSA) allows relating empirical multivariate measures of brain activity to theoretical models (Kriegeskorte et al., 2008). We performed RSA using voltage values on a subject-by-subject basis. Prior to the analyses data were normalized by z-scoring the values across all trials, regardless of the condition. We then constructed empirical Representational Dissimilarity Matrices (RDMs) every three time points, which measure the geometrical distances between all experimental conditions (see MVPA section) and, finally, estimated the relationship between empirical RDMs and theoretical models.

RDMs were built with data from eight conditions, yielding 8x8 symmetrical matrices. These conditions were all the possible combinations between the design variables: *cue prediction* (faces and names); *cue shape* (shapes 1 and 2 for faces, shapes 3 and 4 for names) and *block* (attention and expectation). We employed a Linear Discriminant Contrast (LDC, also known as Crossvalidated Mahalanobis Distance, see Walther et al., 2016) as measure of distance between conditions, for the following reasons: it is a continuous measure, so it is highly reliable, informative, and lacks a ceiling effect; it includes a crossvalidation loop, which makes it less prone to biases; and it is centered around 0 when the true distance is 0 and therefore it is easier to interpret and more generalizable (Nili et al., 2014). We calculated LDC as described in (Bueno & Cravo, 2021). For every time point and each pair of conditions, we calculated the mean of each channel. This was done in two different datasets (train and test), to perform two-fold crossvalidation. We used two-fold crossvalidation to minimize computational costs, while avoiding biases due to random noise (Walther et al., 2016). The distance between the two conditions in the two folds was multiplied by the pseudo inverse covariance matrix between the residuals of the first and the second conditions in the training set, and the distance values were then averaged across the two folds.

Theoretical RDMs (Figure 3.3A) were built based on the expected distances (assigning values of 0 or 1) between conditions, according to different hypotheses. We built three such model RDMs based on: (1) cue shape (increased similarity between cues with same shape, regardless of the block and predicted target); (2) category (increased similarity between cues that predict the same stimulus category); and (3) block (increased similarity between cues belonging to the same block). The next step was to estimate the share of variance that each of these three model RDMs explained. To do so, at each time point we fitted a linear regression with the model RDMs as regressors and the empirical matrix as dependent variable. As a result, we obtained a *t*-value for each model, time point and participant that explained a significant *unique* portion of the variance, above that explained by the other regressors.

We used a non-parametric cluster-based permutation method to infer statistical significance at the group level, against empirical chance levels. First, at the single-subject level we randomly permuted the labels of the theoretical matrices. These permuted matrices were used as independent variables on a linear regression, which was repeated 100 times per participant. This gave 100 chance level *t*-values per participant and model. Then, one *t*-value for each model was randomly drawn per subject and the selected values were averaged. This was repeated 10^5 times to obtain 10^5 permuted group *t*-values. For each time point, the empirical chance distribution was estimated. As expected, this yielded a *t*-value distribution centered around 0. The above and below thresholds were estimated so that they included the 99% of the distribution. Groups of consecutive time points with values outside the previously calculated thresholds were measured generating the null distribution of cluster sizes. Finally, to further ensure correction for multiple comparisons, we used a False Discovery Rate (FDR) approach to determine the smallest cluster size deemed significant for $\alpha = 0.01$ (López-García et al., 2022; Stelzer et al., 2013).

Besides, to estimate the topography of each model we repeated the analyses for each channel separately, estimating the empirical matrix using all time points. That is, at each channel we fitted a linear regression with the theoretical models as regressors.

Time resolved Multivariate Pattern Analysis (MVPA)

We used time-resolved MVPA (Grootswagers, Wardel, et al., 2017) to study stimulus category-specific preparation by classifying faces vs. names. Note that we did not intend to directly classify attention vs. expectation, as this approach would be biased by existing block differences between these conditions. Instead, classifiers were trained and tested to differentiate between faces and names within blocks, and comparisons between attention and expectation were always performed with this base contrast. Also, classifications were done *before* stimulus presentation with cue-locked EEG, where faces and names were anticipated as relevant or probable by the cue (targets were not relevant for this work and thus not analyzed). The steps of the classification were equal for all analyses unless otherwise specified and were performed with voltage values. Classification was performed using MVPAlab (López-García et al., 2022) running on MATLAB. To maximize observations while reducing computational costs, we performed an MVPA classification every three time points. This way features used for classification were trials by channels matrices of raw voltage in single time points $t_{n.}$

We applied two strategies to increase signal-to-noise ratio. First, the trials used were the result of averaging sets of three trials (randomly selected) in each condition (Grootswagers, Wardel, et al., 2017). Then, we employed a smoothing method, based on a moving average filter with a length of 3 time points. For every t_n the features of the previous and the following data points were averaged, so that $t_n = (t_{n-1} + t_n + t_{n+1})/3$. Then, classes were balanced by subsampling the class with more trials so that the number of trials from the two conditions fed to the algorithm stayed the same (Grootswagers, Wardel, et al., 2017). A five-fold stratified cross-validation loop was implemented (King et al., 2013), which ensures that the proportion of each class stays balanced across folds, thus increasing the classification applicability to unknown data. Data were divided into five parts, which are enough to obtain unbiased results with a relatively small computational cost (Grootswagers, Wardel, et al., 2017; Varoquaux, 2018). The classification algorithm was trained with the first four divisions (training set) and tested on the remaining one (test set). This was repeated five times with the different sets. To improve the performance of the classifiers and the generalizability of multivariate analyses results (Singh & Singh, 2020), we normalized the data. Normalization was carried out within the cross-validation loop (King & Dehaene, 2014). Within each fold, we calculated the mean and standard deviation of each electrode across the training trials. The train set (X_{train}) and testing set (X_{test}) were normalized as follows:

$$X_{train} = \frac{X_{train} - \mu_{train}}{\sigma_{train}} \qquad X_{test} = \frac{X_{test} - \mu_{train}}{\sigma_{train}}$$

Where μ_{train} and σ_{train} are the mean and standard deviation of the training set.

We used Linear Discriminant Analysis as classification algorithm, given its good fit to typical EEG variability and higher sensitivity than similar methods (e.g. Support Vector Machines, see Grootswagers, Wardel, et al., 2017; Kerrén et al., 2018). Time ranges were from -100 ms to 1550 ms. Classification results were estimated with an empirical receiver operative curve (ROC) analysis and reported as the area under the curve (AUC). This method works as an estimate of the true positive rate as a function of the false positive rate. The AUC is a non-parametric criterion-free method, so it does not involve assumptions about the true distribution of the data (King & Dehaene, 2014). Besides, it is less susceptible to systematic biases and it is especially sensitive to two-class differences. AUC results can be interpreted similarly to classification accuracy, with 0.5 indicating equal probability of true positives and false positives and 1 accounting for perfect discriminability between classes (King et al., 2013).

To estimate statistical significance, we again used cluster-based permutation analyses. In this case, the labels of each trial were randomly permuted. This was repeated 100 times per participant, generating chance level results. After following the same process as in the RSA section (using AUC instead of *t*-values) we ended up with a distribution centered around chance levels (0.5).

Temporal Generalization Analyses

To characterize the changes of the signal throughout the temporal window, we employed a temporal generalization approach (King & Dehaene, 2017). On each time point we trained a classifier following the process described above. Then, we tested it on all time points of the preparation time window. This rendered a Temporal Generalization Matrix (TGM) representing the AUC values for each train-test pair. Statistical significance was then calculated for each TGM following the same rationale as with time-resolved analyses. The only difference was setting the minimum statistically significant threshold for cluster sizes to $p<0.001$, to avoid small clusters.

Category-specific anticipation within attention and expectation contexts

The process described was applied to the cue-locked EEG separately for attention and expectation blocks, training classifiers to tell apart data from trials in which the cue anticipated faces vs. names. In a first approximation, we used trials of all cue shapes in each category (Figure 3.1C). Next, to ensure that classification results were not biased by perceptual differences between the geometrical shapes used as cues, we implemented a classification approach across cues. We first trained the classifier with data from only one pair of cues (e.g., classifying between squares predicting faces and diamonds anticipating names) and then tested it on the other pair (e.g., testing with circles that predicted faces and drops that predicted names). Because of the design, the selected training and testing pairs of cues only appeared together in half of the blocks. This cross-classification ensures that only the common differences between the two classified pairs will be decoded from the results (Kaplan et al., 2015), thus removing perceptual confounds. We averaged the results of both directions and their permutation maps to obtain a greater signal-to-noise ratio and to reduce biases due to the classification of specific perceptual features. The averaged results were then fed to the same statistical algorithm used previously to obtain cluster-based thresholds of statistical significance.

Once the results of these classifications were obtained, we compared the scores in attention vs. expectation blocks. To do so, we subtracted the empirical results of the two conditions (attention – expectation). Similarly, we subtracted the results of one of the 100 permutated chance level accuracy scores obtained during the cluster analysis in one condition from one from the other. We used the same cluster-based permutation implementation described above to evaluate differences in the two directions (attention > or < expectation). In this analysis the permuted distribution is centered around zero. Since we (arbitrarily) subtracted attention - expectation, positive values indicated greater results for attention, while negative values indicated greater results for expectation.

Cross-classification between attention and expectation

To estimate the degree to which patterns of brain activity are shared for preparation across attention and expectation, we employed a cross-classification approach. We trained a classifier with data from one condition and then tested it on the other. We first did this using data from all trials in each condition, incorporating all four types of cues. Again, to rule out perceptual confounds we repeated the analyses following the same rationale described in the previous section (see Figure 3.1C for a summary of the condition selection strategies). In addition, we performed a control analysis to ascertain that cross-classification between attention and expectation blocks was feasible. Here we trained and tested using cues with the same physical form (e.g. train circle vs. square in attention, test circle vs. square in expectation) to observe whether the classifiers could extract the physical patterns of the cues even across overall changes in block demands.

Canonical Template Tracking

Finally, we compared the sustained patterns that arose during the preparatory interval with the actual perception of face and name stimuli. To do so, we obtained Canonical Template Patterns (CTP, González-García et al., 2021; Palenciano et al., 2022; Wimber et al., 2015) of brain activity generated by faces and names in the independent localizer blocks. First, we performed an MVPA analysis in localizer trials following the same process as in previous analysis. Then, we selected the time window where the classifiers locked to the localizer target stimuli in the localizer had the highest AUC (i.e., when activity patterns were more dissimilar) across participants (Supplementary Figure 3.1), which was 100 – 300 ms after stimulus onset. Then, and separately for face and name localizer trials, we averaged the raw information of every time-point and trial in the selected window for each channel and category. This resulted in a vector of 64 channel activity values for faces and another one for names in each participant. Next, these CTP of faces and names were used as regressors in a linear regression where the dependent variable was the raw channel activity for each channel and condition during every time-point in the preparation or target window of the main experimental task. This rendered two *t*-values per time-point that accounted for the variance explained for each CTP (faces and names). We did this analysis separately in attention and expectation, for cues predicting faces and names and for face and name targets. To estimate statistical significance, we used the same cluster-based permutation analysis described above. Briefly, before averaging localizer data to create the templates for faces and names, we randomly permuted the trials of both conditions. This was repeated 100 times for each stimulus. Then, the randomly permuted CTPs were used as regressors, which gave 100 *t*-values for each participant and template. Then, the process was identical to the one we used in the RSA analysis. To estimate significance when comparing the results for attention and expectation, we employed the same method used to compare MVPA results for both conditions.

Importantly, we used localizer data only from the targets, even though probabilistic cues were also included in these blocks. The localizer only included probabilistic (not relevance) cues, so it would not give equal insight into both relevance and probability contexts in the main task. Also, we considered that overall task demands between the localizer and the main experiment were too large to interpret unambiguously a potential lack of generalization of anticipation between these two contexts. Furthermore, the localizer was not designed to elicit reliable preparatory neural activity, which was hence not tested.

3.4. Results

3.4.1. Behavioral

Forty-eight participants completed a cue-target paradigm where, depending on the block, cues carried information about either the *relevance* (attention) or the *probability* (expectation) of upcoming face or word targets (Figure 3.1A). Analysis of participants' behavioral results showed that preparation for the incoming target stimuli affected performance. A three-way repeated measures ANOVA on reaction times (RTs) showed main effects of Block ($F_{47,1}$=56.45, p<0.001, ηp2 = 0.54), Cueing ($F_{47,1}$=5.59, p=0.022, ηp2 = 0.11) and Category ($F_{47,1}$=50.52, p<0.001, ηp2 = 0.52). Overall, responses were faster in attention (M = 569 ms, SD = 0.01) than expectation (M = 598 ms, SD = 0.01) blocks. Cueing affected RTs differently depending on the block (Block*Cueing,

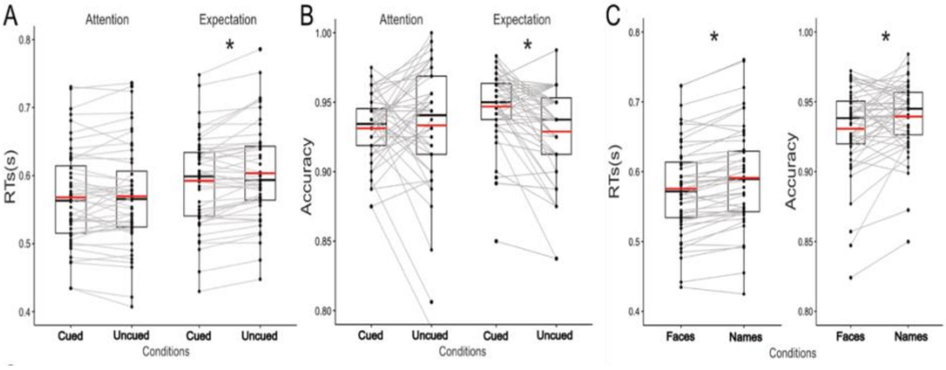


Figure 3.2. Behavioral results. Participants responded with a yes/no key press to the question asked at the beginning of each block. Note that in attention blocks, uncued trials were always responded with a "no" response. (A) Reaction times (in seconds) in attention and expectation blocks, for cued and uncued trials. (B) Accuracy in attention and expectation blocks, for cued and uncued trials. (C) RTs and accuracy for face and name targets, averaged across conditions. Dots represent individual subject (n=48) scores per experimental condition. Grey lines connect each participant's score in the two conditions of each block. The horizontal black line inside boxes represents the median, the horizontal red line represents the mean, while the limits of the box indicate the first and third quartile. Note that due to the high SD, the mean and the median may differ slightly. Whiskers indicate the 1.5 inter quartile range for the upper and lower quartiles.

$F_{47,1}$=9.07, p=0.004, ηp2 = 0.16; Figure 3.2A). Post-hoc tests showed no effect of Cueing for attention (t<1) whereas there was an effect for expectation ($t_{47,1}$=3.6, p<0.001, Cohen's d = 0.52). Expected (cued) targets elicited faster responses on average (M = 592 ms, SD = 0.06) than unexpected (uncued) targets (M = 603 ms, SD = 0.07). Category also induced significant differences. Post-hoc tests showed faster responses for faces (M = 576 ms, SD = 0.7) than for names (M = 591 ms, SD = 0.75), $t_{47,1}$=7.51, p<0.001, Cohen's d = 1.08; Figure 3.2C.

The ANOVA on behavioral accuracy showed results in the same direction albeit less prominent. Both Cueing ($F_{47,1}$=4.33, p=0.043, ηp2 = 0.01) and Category

($F_{47,1}$=7.04, p=0.011, $\eta p2$ = 0.011) were significant, and there was no main effect of Block (*p*>0.05), with both conditions presenting high accuracy overall (attention: M = 0.92, SD = 0.07; expectation: M = 0.93, SD = 0.07). Again, we found that cueing affected each block differently (Block*Validity, $F_{47,1}$=9.67, p=0.003, $\eta p2$ = 0.15; Figure 3.2B). Although there were no differences in attention (t<1), there was an effect of expectations ($t_{47,1}$=3.59, p=0.003, Cohen's d = 0.52), with cued trials eliciting more accurate responses (M = 0.947, SD = 0.026) than uncued ones (0.929, SD = 0.035). Post-hoc tests showed less accurate responses to faces (M = 0.93, SD = 0.4) than to names (M = 0.94, SD = 0.37; Figure 3.2C), $t_{47,1}$=2.65, p=0.011, Cohen's d = 0.38. Overall, these results indicate that, as instructed, cues were used effectively and also differently across blocks.

3.4.2. Time-resolved profile of preparation

Our first aim was to assess the emergence of specific coding patterns linked to different information content during preparation in attention vs. expectation contexts. The perceptual features of the cue should be processed, and their contextual meaning extracted to anticipate general target-category information. This anticipated target category should be activated and maintained in working memory to later provide an efficient response. To assess the contribution of each of these sources of information to preparatory activity we employed model-based RSA (Kriegeskorte, 2008).

Results of the multiple regression for the three RSA models built (Cue shape, Category and Block, see Methods section and Figure 3.3A) revealed the unique variance explained by each of the factors entered in the analysis. A peak of the Cue shape model appeared first, at 160 ms after cue presentation, and decayed fast afterwards (Figure 3.3B, red line). In contrast, the coding of the specific incoming target category increased progressively along the interval, reaching its peak right before the presentation of the actual target (Figure 3B, green line). In addition, the variance explained by the blocks reached its peak around

500 ms after cue onset, slowly decayed along the interval, and fell below significance levels 100 ms before target onset (Figure 3.3B, blue line).

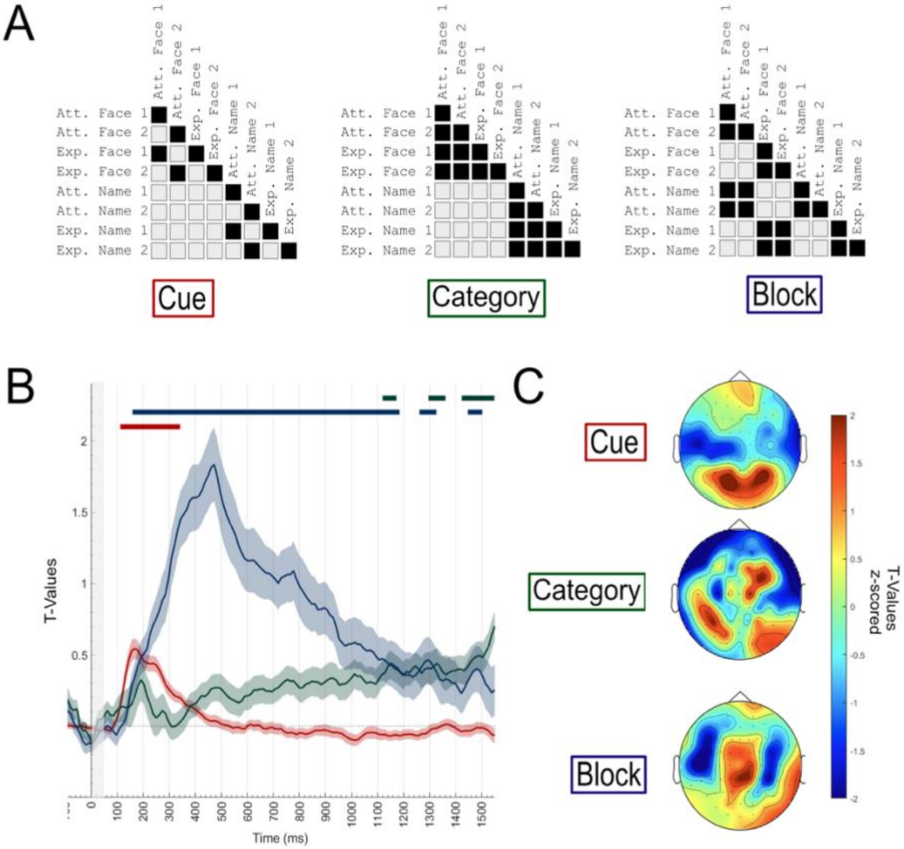


Figure 3.3. RSA theoretical matrices and results. In (A), black squares represent high similarity between conditions and grey squares represent low similarity. In (B), each line depicts t-values from a linear multiple regression fit for the three theoretical dissimilarity models, showing the unique share of variance explained for each of the factors. The colored straight lines above indicate significance clusters after a cluster-based permutation analysis. Grey shading indicates cue presence onscreen. The cue model (red) indicates dissimilarity between cue shapes, regardless of the block or prediction; the category model (green) indicates similarity between cues predicting faces or names; the Block model (blue) shows overall preparatory differences between attention and expectation. (C) Topographies of each model after using time points as features and repeating the analysis on each EEG channel.

3.4.3. Time-resolved classification shows increasing AUC throughout the preparation interval.

We then studied the coding of the anticipated specific information during the preparation interval, separately for attention and expectation. Importantly, we did not compare cued vs. uncued targets, but cues that predicted (relevant or probable, depending on the block) face vs. word stimuli within the same block. Using the two cues of each category together for training and testing returned a classification weighted on the perceptual features of the cues (see Supplementary Figure 3.2 for a detailed description of the result), similar to the Cue model in Figure 3.3 (red line). Afterwards, we employed a cross-classification analysis between different cues to avoid cue perceptual confounds in the classification (Figure 3.4). Since two differently shaped cues coded for each type of category, we trained the classifier in one pair and tested it on the other, repeated the process in the opposite direction and averaged both. Similar to the Category model in Figure 3.3, time resolved cross-classification showed that the accuracy in the decoding of the anticipated category increased as the target onset approached. We employed the

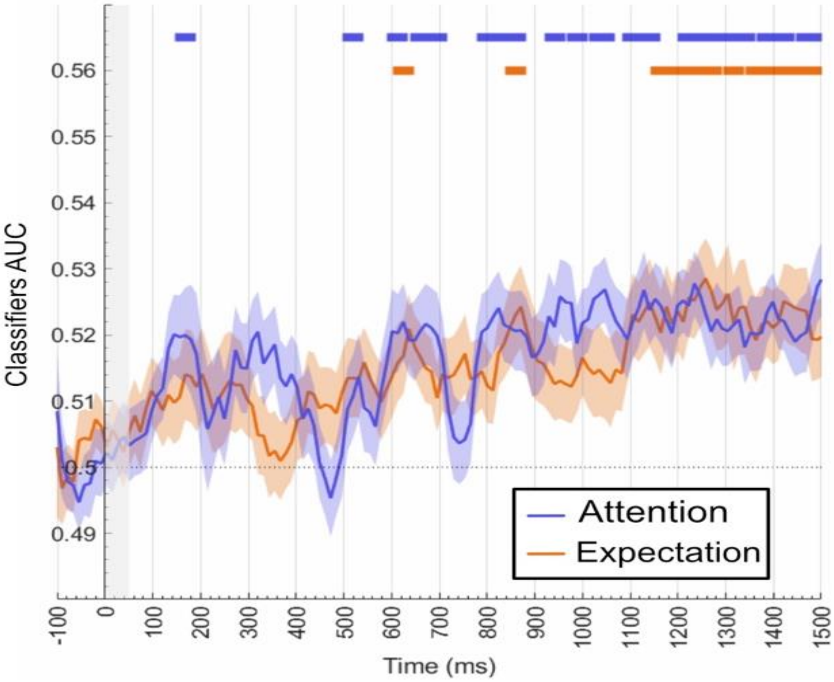


Figure 3.2. Raw voltage time-resolved cue decoding results. Result of the time-resolved classification of the category (faces vs. names) to be selected (blue) vs. expected (orange). Horizontal colored lines indicate statistical significance against chance within each block. Grey shading indicates cue presence onscreen. Figure shows the result of training and testing as described in Figure 3.1C (dotted line). For the results using all the cues together for training and testing see Supplementary Figure 3.2.

subtraction approach described in the Methods section, but we found no differences in classification between attention and expectation (all ps>0.05). We repeated a regression analysis to predict the cross-classification result of each participant based on the average time-resolved result (*t*-values) of each RSA model. As expected, the category model (Figure 3.3, M = 0.35, SD = 2.37) explained the results better than any other. *T*-values were nominally higher than the cue model (M = 1.04, SD = 1.42) and significantly higher than the block model ($p<0.01$).

3.4.4. Relevance- and probability-driven preactivations are stable

While the previous results provide an initial characterization of neural coding of specific category information during the preparation interval, they do not allow to explore the extent to which relevant representations are stable during this time window, since significant decoding on different time points could be driven in principle by different mechanisms. To investigate this, we employed a cross-time decoding approach (King & Dehaene, 2014) to compare different patterns of brain activity across the preparation interval. For this, we trained a classifier in one time point and then tested it on all the points of the interval (Figure 3.5A, B). Results showed clear signs of generalization during the preparation interval in both attention and expectation conditions (black outlines). The clusters of activity grew increasingly larger up until the target's onset, indicating that the underlying patterns remain relatively stable during the preparation period. We then compared the results of both analyses following the same rationale as above: we subtracted the result matrices for both conditions and performed a one-side *t*-test against 0. This analysis did not yield any significant results, indicating that the accuracy of anticipatory category decoding was not different between conditions. Altogether, these results suggest that preparation in both attention and expectation leads to a similar level of discriminability of the anticipated (relevant or probable) category. However, this analysis is agnostic regarding potential similarities in how anticipated relevant vs. probable information is coded, as different underlying mechanisms could lead to similar accuracy results.

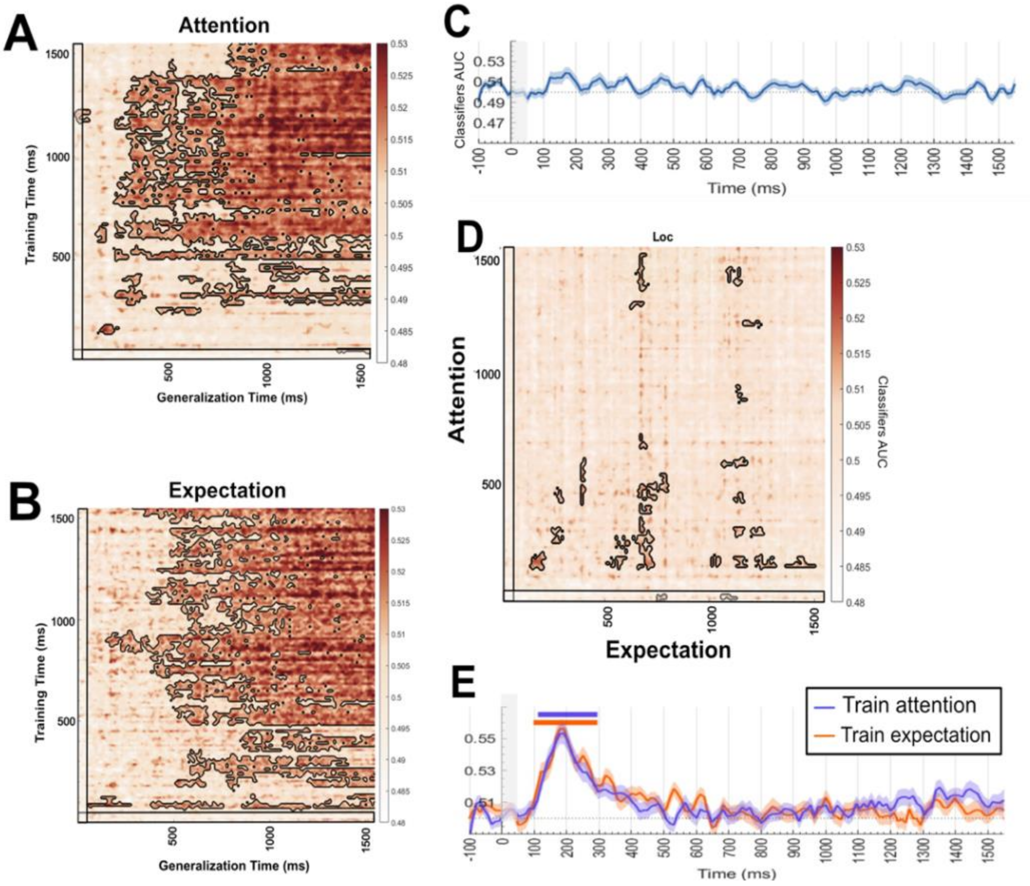


Figure 3.3. Raw voltage temporal generalization results. (A-B) Temporal generalization results using the cross-cue decoding scheme, within (A) attention and (B) expectation conditions. Black outlines indicate statistical significance against chance. Black boxes indicate cue presence on screen. (C-D) Cross-condition classification results. Classification using the scheme described in Fig 1C, depicting the average of the two train and test directions. (C) Time-resolved classification of anticipatory information coding after training and testing the classifier in different conditions yielded no significant decoding. (D) Temporal generalization matrix. For visualization purposes, we averaged training and testing in the two directions (train attention and train expectation). The horizontal axis shows the times for training and testing in expectation, and the vertical axis that of training and testing in attention. That is, we inverted the axis of one of the directions before averaging. We chose this rationale because activity that is better represented in only one of the conditions should appear only on one side of the diagonal (e.g., above) when training and the opposite when testing (e.g., below). Averaging this way, we can show information that would not be visible doing a standard average. The non-averaged results for both directions can be found in the Supplementary Section (Supplementary Figure1). (E) A control cross-classification between attention and expectation (Orange: Train attention – Test expectation. Purple: Train expectation – Test attention) employing the same pairs of cues to train and test the classifier (e.g. train with circles in attention, test with circles in expectation) shows that classification across blocks is feasible. However, whereas the perceptual features of the cues generalize across attention and expectation (significance is marked with colored bars), the specific preactivation of contents before target onset does not.

3.4.5. Attention and expectation induce distinct patterns of preparatory activity

To examine the mechanisms supporting the classification results, we used multivariate cross-classification by training the classifier in one condition and then testing it on the other. Similar patterns of brain activity should generalize between conditions, while differences should provide chance-level classification. The number of observations employed for this, as well as the inter-block temporal distance, were also matched. The averaged results (training in attention and training in expectation) for raw voltage cross-classification are shown in Figures 3.5C, D (see Supplementary Figure 3.3 for the results split by train and test direction). Common coding between anticipating relevance vs. probability of stimulus categories was scarce. Small significant clusters appear scattered through the temporal generalization interval. This result complements the previous analysis by suggesting that different neural mechanisms support the classification results. In addition, the control analyses performed using the same physical cues to cross-classify between attention and expectation (see Figure 3.5E) show that generalization across blocks is feasible, given that the classifier shows significant above-chance performance to predict the perceptual shape of cues across attention and expectation conditions.

3.4.6. Tracking perceptual patterns of brain activity

We have shown that although attention and expectation lead to similar degrees of anticipatory classification, their underlying neural patterns are partially different. We hypothesized that such differences could arise due to the extent to which anticipatory representations function as perceptual reinstatements of the prepared target categories. In this case, the preparatory patterns that allow to classify anticipated faces and names should be similar to those triggered during actual stimulus perception. To assess such reinstatement of target representations during preparation, we obtained CTPs (Palenciano et al., 2023) of brain activity associated with faces and names (as described in the Methods section). We then fit a linear regression with the two CTPs as regressors to measure the extent to which these explained variance in preparatory activity across conditions, in each time point (see Figure 3.6).

First, to complement the behavioral results regarding face and name processing, we applied the CTP to EEG activity locked to target onset in attention and expectation (Supplementary Figure 3.4). The face CTP explained better the perception of target faces, while same was true for the name pattern on name perception. Moreover, face processing was more pronounced, but decreased more rapidly. Name processing, on the other hand, reached a lower peak but it decreased more slowly over time. As expected, these results further suggest that faces and names are processed differently.

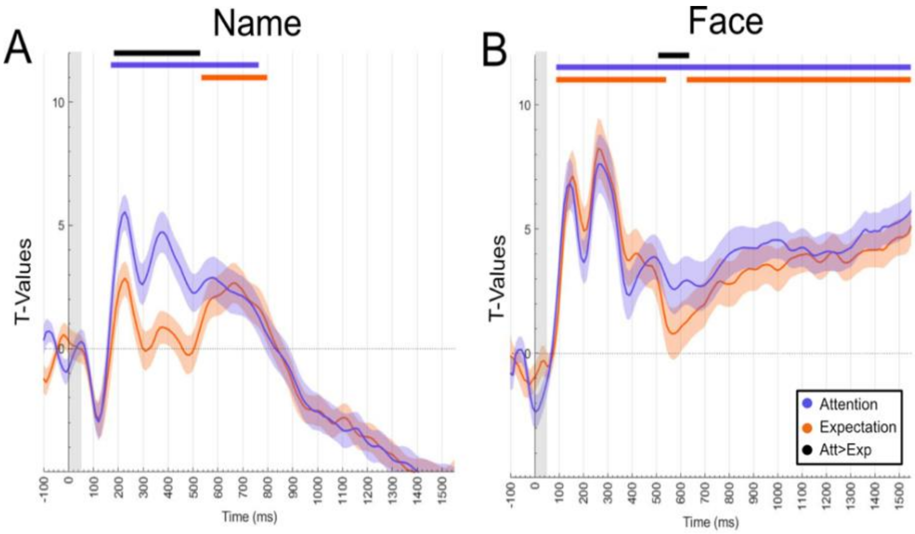


Figure 3.6. Canonical Template Tracking results. Grey shading represents cue on screen. (A-B) CTP results for separate cues for attention and expectation. Average of face and name cues (A) Name CTP model for cues. (B) Face CTP for cues. Horizontal lines depict significance. Black horizontal lines indicate points where the results for attention were significantly higher than for expectation. There were no clusters were expectation showed higher reinstatement.

Then, we compared how these canonical templates overlapped with preparatory activity in attention vs. expectation contexts. Since both CTPs similarly explained face and name predicting cues, here we show the average result of the two types of cues separately in attention and expectation blocks. Note that whereas such averaging prevents drawing any conclusion regarding selective preparation, it allows contrasting the extent of overall perceptual reinstatement across relevance vs. probability anticipation. Crucially, both perceptual templates were better predictors of category anticipation in the attention compared to the expectation condition (see Figure 3.6). Overall, these results indicate that regardless of the intrinsic differences in face and name processing, there is an overall stronger perceptual reinstatement for attention than for expectation. Note, importantly, that this stronger coding cannot be explained by overall increased difficulty, as responses were faster in attention compared to expectation blocks (see Behavioral Results section).

3.5. Discussion

In the present work we studied and contrasted, for the first time, the top-down neural mechanisms engaged during category-specific preparation in two major information-processing contexts: selective attention and perceptual expectations. Our results reveal that the anticipated stimulus category is active during preparation, regardless of the specific demands. Crucially, we demonstrate that relevance and probability influence the preparatory patterns of brain activity in a unique manner, potentially via differences in the perceptual reinstatement of anticipated information. Overall, these results have important implications for theoretical models that explain how the brain anticipates relevant vs. probable forthcoming information.

Our behavioral results show that the paradigm manipulated attention (relevance) and expectation (probability) in an effective manner as expected (Baldauf & Desimone, 2014; Kok et al., 2017). Cued trials in attention required further processing than those uncued, as participants were required to evaluate the sex/gender of the cued stimulus category and skip this judgment in uncued ones (Baldauf & Desimone, 2014; Saenz et al., 2002; Summerfield et al., 2006; Womelsdorf et al., 2006). Importantly, for attention the probability of either category was equal (50%), keeping expectation constant for relevant and irrelevant trials. The high accuracy in attention blocks indicates that the cues were used effectively to select when needed and to respond. On the other hand, both expected and unexpected trials required selecting the forthcoming target, and thus equated relevance while leading to predicted and unpredicted stimulus categories. In line with validity effects repeatedly reported in the literature in expectation contexts, the efficiency of behavior increased in cued trials (de Lange et al., 2018; Sy et al., 2014). It may be important to note that our results are at odds with the notion that expected stimuli gain higher task-relevance and thus receive more attention that unexpected ones (see Rungratsameetaweemana & Serences, 2019), as this interpretation would lead to expect similar underlying anticipatory patterns for our attention and expectation manipulations.

We employed RSA (Carlson et al., 2019; Kriegeskorte, 2008) combined with multiple regression to study the temporal profile of information coding throughout the preparatory interval. This showed that anticipatory cues triggered several time-overlapping, yet distinguishable, effects in time (Figure 3.3B) and spatial arrangement (Figure 3.3C). The perceptual characteristics of the cue were coded early in the time interval. Immediately after, and during the majority of the interval and until around 100 ms before target onset, mechanisms pertaining to contexts capitalizing on either selective attention or expectations were deployed, as revealed by the Block model temporal profile. Evidence of anticipatory coding of different task sets has been found before (González-García et al., 2021; Hebart et al., 2018; Palenciano et al., 2018). Interestingly, this result cannot be explained by motor preparation, since cues did not predict motor responses; instead, it is most likely due to the control context in each scenario. Finally, the variance explained uniquely by the Category model on anticipatory activity increased steadily and peaked before target onset. This pattern of emerging, ramping-up category representations could reflect two non-exclusive anticipatory mechanisms. Once the meaning of the cue is extracted and stored in working memory, the actual anticipation of a stimulus could be facilitated by top-down preactivation of perceptual regions (Auksztulewicz & Friston, 2016; Trapp et al., 2016). Simultaneously, the effect of temporal expectations (as the preparatory interval was fixed) could induce an increasing preactivation of perceptual regions as the stimulus onset approaches (Jin et al., 2020; Rohenkohl et al., 2012). In either case, these results suggest that preparation engages a series of mechanisms that act both sequentially and in parallel, including the processing of bottom-up signals that are transformed into top-down category anticipation in a context of relevance or probability.

The block model in the RSA suggested that attention and expectation led to distinguishable coding patterns. One mechanistic explanation for this result is that the patterns for the anticipated category are more robustly coded in one condition over the other. Hence, we employed an MVPA approach (Grootswagers, Wardel, et al., 2017; Haxby et al., 2014) to classify the prepared category. Crucially, classifiers did not directly compare attention vs. expectation blocks (a discrimination that would be biased by differential block characteristics) nor cued vs. uncued targets (which would mix reactions to matched vs. mismatched predictions). Instead, classifiers were trained and tested to differentiate between cues that predicted (relevant or probable) faces vs. names, within each block. Results matched the RSA model, pointing to a robust effect of category anticipation. Furthermore, this prediction increased towards the end of the interval. Again, this resonates with literature on temporal anticipation (Barbosa et al., 2020; Jin et al., 2020; Ruz & Nobre, 2008). Moreover, prominent theories of attention (Mongillo et al., 2008; Trübutschek et al., 2017) pose that information can be held in WM by strengthening the synaptic weights between neurons, allowing for shifts (in this case, steady increases) in decoding results during the delay period without losing information about the maintained stimulus. The results revealed no significant differences in AUC between the conditions in either time-resolved or temporal generalization matrices, and suggest that anticipated targets are coded with similar fidelity during expectation and attention. Importantly, these results were obtained cross-classifying different pairs of cues, thus avoiding perceptual confounds. However, a similar degree of accuracy classification does not warrant identical underlying mechanisms. To test this, we performed multivariate cross-classification between attention and expectation. Surprisingly, despite well above-chance decoding within each type of block, there was little cross-classification between conditions. Importantly, classification scores were similar when training and testing in different blocks of the same condition (see Supplementary Figure 3.5). In addition, a control analysis using the same cues showed that generalization across blocks was possible and not prevented by the differential block context. Classifiers trained and tested in different blocks were able to discriminate the perceptual nature of the cues, whereas the anticipated category cross-classification was absent (see Figure 3.5). Overall, this set of results suggests that whereas attention and expectation both lead to anticipatory category representations, their top-down mechanisms are partially different, providing further support to the dissociation of relevance and probability, reported by previous studies, during target processing (Summerfield & Egner, 2009; Wyart et al., 2012).

One possible explanation for the differences found is the degree of preparatory perceptual reinstatement (Kerrén et al., 2018; Kilner et al., 2007; Muckli et al., 2015; Rose et al., 2016; Smith & Muckli, 2010; Vetter et al., 2014) in attention and expectation. We employed a Canonical Template Tracking procedure (González-García et al., 2021; Wimber et al., 2015) to compare perceptual reinstatement in attention and expectation. We obtained canonical representations of the two target categories (faces and names) from an independent localizer, and then estimated the variance explained by these perceptual patterns during the anticipatory window. First, we applied the extracted canonical templates to the actual target processing, with results supporting that faces are processed differently from words. Importantly, when applied to anticipatory activity, we found that the canonical templates explained preparatory variance equally well for both predicted categories, but the reinstatement was significantly higher in attention than in expectation. Note that this cannot be due to a higher difficulty of the attention blocks, as RT showed the opposite pattern (faster responses in attention). Instead, this higher reinstatement may be due to attention directing more resources to activate perceptual codes in anticipation. It is unclear, however, why this analysis did not provide evidence of category-specific reinstatement, although it could be related to the mixture of perceptual activity caused by the cues themselves. It is possible that the physical shape of faces and words has differential overlap with the overall shape of the cues employed in the main task, which could have added additional variance to the overall analysis. This could have mixed with the anticipated category information and generated the lack of specificity. Relatedly, the patterns captured by our CTP may be different from those involved in categorical specific anticipation, which may have happened at a different abstraction level. Note that the large overall task differences between the localizer and the main task prevent the use of localizer cues to predict preparation during the main task. Further research employing different approaches to measure reinstatement should be conducted to clarify this matter.

How attention and expectation affect perception is an ongoing debate. Several studies have used paradigms that combine both processes to study how they interact. Attention has been suggested to sharpen the differences between expected and unexpected stimuli (Jiang et al., 2013), possibly changing the oscillatory profile of relevant categories (Auksztulewicz et al., 2017) while reversing repetition suppression (Kok, Rahnev, Jehee, Lau, & De Lange, 2012). Although it has been suggested that attention acts from early processing stages, results so far are not conclusive (Alilović et al., 2019). Relatedly, attention boosts signal-to-noise processing by suppressing noise, while expectation increases baseline activity in perceptual regions (Wyart et al., 2011) see also (Gordon et al., 2019; Rungratsameetaweemana & Serences, 2019). Predictive coding accounts propose that attention increases prediction error of selected stimuli by suppressing noise of unattended categories, while expectation increases global sensitivity through prediction signals. Crucially, our results extend this literature by showing anticipatory differences between conditions, which cannot be accounted for by prediction error differences as targets have not been processed yet. Speculatively, attention could bias anticipatory neuronal sensitivity by increasing perceptual differences between relevant categories, coding these changes at least partially in the gamma band perhaps through anticipatory biasing of error processing units.

Expectation could increase sensitivity to probable categories by increasing excitability of, perhaps, perceptual units.

Although overall our results are a crucial first step to characterize mechanisms across relevance and probability anticipation, they should be complemented by further studies. Although the results are statistically significant after robust cluster-based correction, accuracy values are lower than those obtained, for example, using target-locked data (e.g., Supplementary Figure 3.1). Importantly, decoding accuracies do not equal effect sizes (Hebart & Baker, 2018). The values we obtained are within typical classification ranges when studying subtle neural patterns (Christophel et al., 2015; Hebart & Baker, 2018; Rose et al., 2016) employing non-invasive human neuroimaging. Future studies may obtain higher accuracies by increasing stimulus repetition and thus reinforcing specific neural traces of the information anticipated.

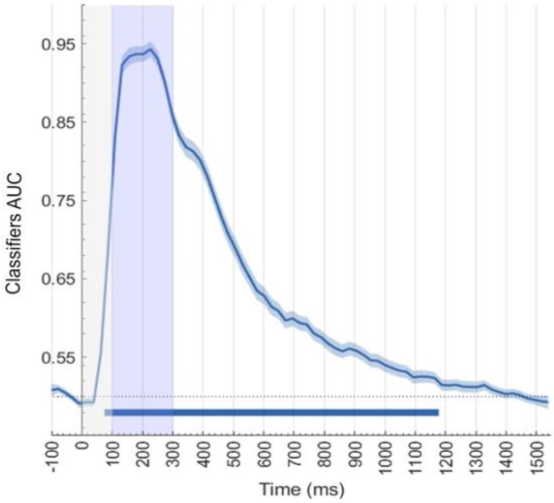
Moreover, it could be argued that the diverging results between the two conditions shown here are not due to differences in stimulus representation for relevance vs. probability, but to different task demands of the two blocks. However, as the goal of the study was not to fully differentiate the neural mechanisms of attention and expectation but to understand how anticipatory information is coded differently in contexts of relevance vs. probability, we did not merely contrast the two blocks directly. Instead, the analyses contrasted anticipated faces vs. words in conditions matched within blocks. In the same line, the arguably potential higher relevance of probable contents could not drive results, as such increased relevance would be equal for predicted probable faces vs. names, and thus the classifier cannot rely on this information. Moreover, and to further show how processes that are similar do generalize across both blocks, we trained the classifier in a pair of cues in one block, and tested it in the same pair for the opposite block (see Figure 3.5E). As expected, significant classification appeared for the cue identity, showing that cross-classification across blocks is feasible. If differential task demands had an overall effect changing the format of coding during anticipation, this should have arguably also altered cue patterns, preventing cross-classification. The fact that classifiers were able to extrapolate, though, adds to the idea that block differences do not account for lack of cross-classification of anticipated relevant vs. probable information content. Admittedly, classifiers can be conceived as black boxes where disentangling which factors are driving the results can be hard. However, the results obtained when employing RSA provide concurring evidence. Here, theoretical matrices (Cue, Category and Block) were used to study how these factors explain unique variance during the preparation window. Results (see Figure 3.3) suggest that Block (task demands) explains most of the variance during an earlier time window than Category, which in turn incrementally explains variance (and with a different topography from that of Block) as target onset approaches. Of note, the timing of this effect is quite similar to the window where classifiers show specific preparation for relevance and probability that does not cross-classify across these contexts. In any case, studies employing improved paradigms should be tested to replicate and validate these results.

Our scope was limited to the temporal domain, and questions arise regarding potential differences between brain regions. Internal predictions have been generally associated with the hippocampus (Aitken & Kok, 2022; de Lange et al., 2018; Hindy et al., 2016; Stachenfeld et al., 2017), which has a location that challenges EEG sensitivity. Further studies should employ more spatially sensitive techniques. Additionally, we focused on category-based preparation of faces and words, to facilitate having the same task across categories (sex/gender judgments). These two types of stimuli, although frequently combined in the literature (Alm et al., 2016; Amado et al., 2018; Dumas, 2015; Rose et al., 2016; Sperling et al., 2003) have a different spatial layout, which may have generated the anticipation of different spatial templates. Although this is not a confound in our task, as such difference is constant, it may have added a spatial component to the preparation. Finally, we focused our design on visual perception. Studies that have compared the effects of attention and expectation have used auditory and visual stimuli showing promising results

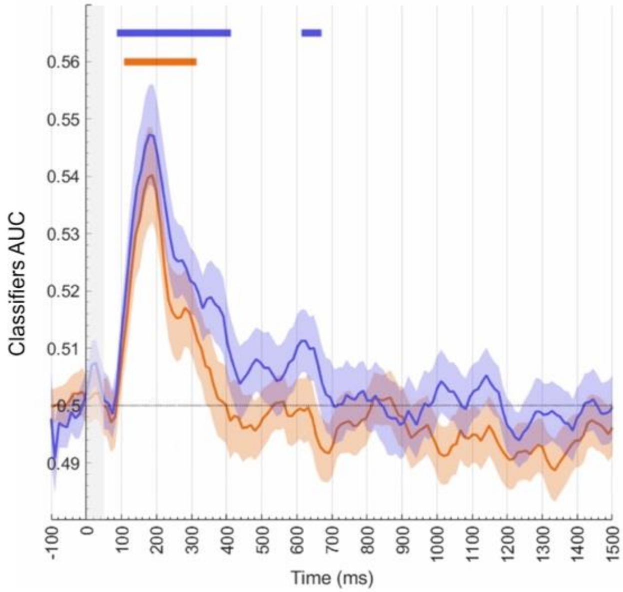
(Jiang et al., 2003). It is likely that the mechanisms found here can be extended to other sensory modalities, but more research is needed to assess this idea.

Altogether, our results show that anticipatory representations are category-specific and also task dependent, and highlight the role of proactive processes that precede stimulus perception. These findings have important implications for current models of brain functioning. We show that instead of single, unitary top-down phenomena, the brain implements distinct modes of content-specific anticipation, which are tailored to task context and involve different levels of overall perceptual reinstatement. Predictive coding and attention models that differentiate these processes during target processing should be extended to the preparatory interval, acknowledging the specificity and complexity of these top-down anticipatory phenomena.

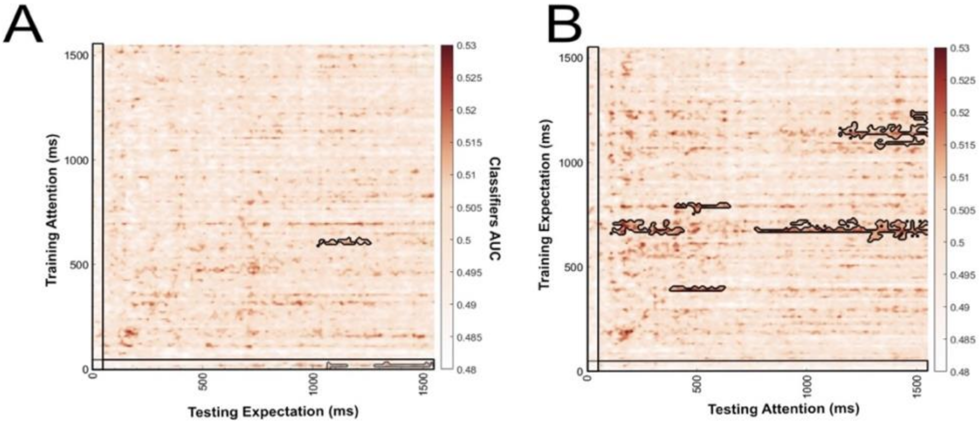
3.6. Supplementary Materials



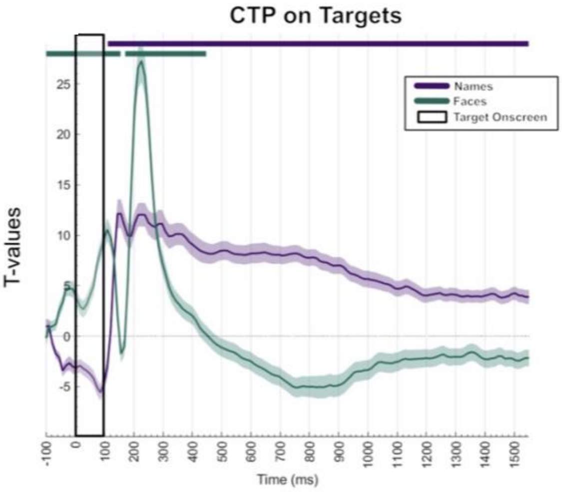
Supplementary Figure 3.1. Localizer's result after classifying faces and names. Grey shading indicates stimuli presentation. Blue shading shows the selected time window to create the two template patterns for faces and names.



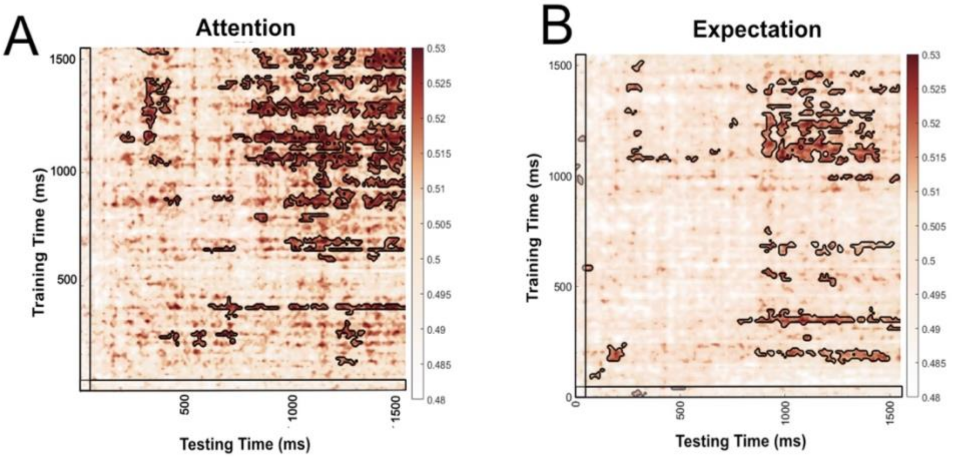
Supplementary Figure 3.2. Classification result using all cues to train and test the classifier. Result of the time-resolved classification of the category (faces vs. names) to be selected (blue) vs. expected (orange). Horizontal colored lines indicate statistical significance against chance within each block. Grey shading indicates cue presence onscreen. Training a linear classifier with the two cues of each condition together to decode stimulus category separately for attention and expectation showed early peaks in classification that quickly vanished. We compared both AUC results by subtracting them and performing a cluster-based permutation analysis, which yielded no significant differences (all ps>0.05). To directly compare MVPA and RSA results, we performed a regression analysis using the RSA results as predictors of each participant's decoding accuracy. We then compared the variance explained by each regressor (i.e. RSA model) in a one-way ANOVA. Post-hoc comparisons revealed that, as expected, the Cue model explained the decoding results depicted in this figure better than any other RSA model (all ps<0.001).



Supplementary Figure 3.3. Cross-classification results by training and testing direction. (A) Train attention-Test expectation. (B) Train expectation-Test attention. Black outline depicts significant clusters.



Supplementary Figure 3.4. Canonical Template Tracking results for targets. Since there were no statistical differences between attention and expectation, the results are averaged. The green line shows the results for how much the face CTP explains face perception. Purple shows the result for the name CTP on name stimuli.



Supplementary Figure 3.5. Classification results only using trials from different blocks (A-B). Note that the patterns of generalization are stable in both conditions. However, the AUC results appear reduced, likely due the loss of power after using less trials to train and test.

Chapter IV
Study II

The content of this chapter is in preparation as Peñalver, J.M.G.; González-García, C.; Palenciano, A.F.; Ruz, M. Anticipating relevant vs. probable content involves dissociable neural mechanisms.

4.1. Summary

Preparation is a top-down phenomenon known to improve performance across different situations. In light of recent electrophysiological findings that suggest that anticipatory neural preactivations linked to preparation are context-specific and not generalize across domains, in the current investigation we sought to extend these findings to spatial grounds by using fMRI. We applied multivariate decoding to data obtained in a paradigm where, in different blocks, cues provided information about the relevance or probability of incoming target stimuli. Results showed that the anticipated stimulus category was pre-activated in both conditions, mostly in different brain regions within the ventral visual cortex and with differential overlap with actual target perception. Crucially, there was little cross-classification across attention and expectation contexts, indicating lack of common neural coding across relevance and probability contexts. Overall, our results stress the specificity of anticipatory neural processing depending on its informative role.

4.2. Introduction

For several decades, investigations in cognitive psychology have primarily focused on the study of behavior while manipulating external variables. These have led to theoretical frameworks that conceptualize cognition predominantly from a reactive standpoint. In recent times, there has been a resurgence of interest in proactive cognition, highlighting the central role of endogenous top-down brain processes. Within this framework, preparation can be conceptualized as an endogenous neural function that takes place prior to perception of the target stimulus and that improves subsequent behavior (Battistoni et al., 2017; González-García et al., 2016). It has been shown in several contexts, including those of selective attention (the ability to select information relevant for behavior; e.g. Nobre & Serences, 2018) and perceptual expectations (the generation of probabilistic predictions based on previous experiences, e.g. Schröger et al., 2015).

A large part of the investigation on attention and expectation has focused on their consequences on target processing (reviewed in Summerfield & Egner, 2009, 2016), which seem to lead to opposing effects on activation levels. For example, studies that use selective cues to highlight the target features of stimuli in non-human primates have found increased firing rates in neurons tuned to relevant features in regions of the ventral visual stream, such as V4 (e.g. Bichot et al., 2005; Burrows et al., 2014; Chelazzi et al., 1998) and the inferior (Chelazzi et al., 1998) and superior (Cohen & Maunsell, 2011) temporal cortex. Studies employing non-invasive neuroimaging in humans show matching results (Kastner et al., 1999; Peelen & Kastner, 2011; Serences et al., 2004). On the other hand, studies of expectation classically show effects of activity decrease (also known as expectation suppression) in visual regions such as V1 and lateral occipital cortex (e.g. Kok et al., 2012; Walsh & McGovern, 2018; see Feuerriegel et al., 2021 for a detailed review). Other relevant line of research has explored the interplay between attention and expectation during target processing, with mixed results. Whereas some studies suggest that selective attention modulates the effects of expectations (Alilović et al., 2019; Auksztulewicz et al., 2017; Jiang et al., 2013; Kok et al., 2012), others do not find such interactions (e.g., Ekman et al., 2017; Yon et al., 2018). Zuanazzi & Noppeney (2019) found activation in the frontoparietal network for selected visual and auditory targets, but no effect for expected visual targets. Meanwhile, (Gordon et al., 2019) linked selective attention to lower levels of the visual hierarchy, and expectations to more complex instances of visual perception. Similar results have been found in other studies (Auksztulewicz et al., 2017; Simon et al., 2018; Wyart et al., 2012). Overall, results suggest separate roles of relevance and probability during target processing (Auksztulewicz & Friston, 2016; Gordon et al., 2019; Simon et al., 2018; Wyart et al., 2012).

At the anticipatory level, separate studies of attention and expectation have shown the preactivation of similar brain regions. Selective attention cues (Battistoni et al., 2017; Nobre & Serences, 2018) preactivate specific shape patterns in visual cortex (Stokes et al., 2009), relevant regions of space (Giesbrecht et al., 2006), and object-selective (M. V. Peelen & Kastner, 2011; Soon et al., 2013) and category (Esterman & Yantis, 2010; González-García et al., 2018) perceptual regions. Relatedly, probabilistic cues lead to the preactivation of specific perceptual templates of oriented gabors (Kok et al., 2017), direction (Ekman et al., 2017), motor patterns (de Lange et al., 2013) or abstract shapes (Hindy et al., 2016). Moreover, in both contexts it has been suggested a particular involvement of the frontoparietal network in the generation of specific anticipatory templates (e.g. González-García et al., 2018; Summerfield et al., 2006; Woolgar et al., 2015), although whether the role of this regions differs between contexts of relevance and probability is not clear.

Altogether, results seem to suggest that attention and expectation involve the preactivation of relevant perceptual mechanisms, but nevertheless generate different consequences on target processing. In turn, this poses the question of the specificity of the computations involved in anticipatory brain activity: is preparation a unified phenomenon or, conversely, does it reflect a plethora of different mechanisms? On this regard, a recent EEG study applying multivariate analyses successfully contrasted selective attention and perceptual expectations, crucially, during the preparatory interval (Peñalver et al., 2023). Researchers employed a paradigm in which, in different blocks, cues provided information about the relevance or probability of face and name stimulus targets. Anticipatory coding of incoming categories was found in both contexts, but the underlying neural codes did not generalize across them. Global condition differences between attention and expectation, as well as ramping-up target decoding were found with temporal precision during the cascade of events that unfold during anticipation. However, electrophysiological results are agnostic regarding the brain regions involved in these processes. Differences between relevance and probability anticipation could be due to separable anticipatory neural codes in the same regions, or (potentially similar) activity in different brain areas. Previous research has found the involvement of both perceptual and frontoparietal (e.g., Stokes et al., 2009, González-García et al., 2017) regions during target anticipation.

Here, we aim to further understand the bases of top-down preparation by using fMRI to study whether the anticipation of specific stimulus categories engages similar or context-dependent brain areas. We adapted the task used in Peñalver et al. (2023) to fMRI, which allows to replicate and consolidate the findings that suggested that attention and expectation elicited different kinds of preparation. Given that anticipatory processing is likely to elicit small changes in average brain activity, we performed a series of decoding and cross-decoding analyses, better suited to study the representational characteristics of preparation and to disentangle different patterns of activity that may be taking place in similar regions. Specifically, first we studied the brain regions that, overall, coded information differently depending on whether preparation was performed in a context of selective attention or of perceptual expectations, by classifying attention and expectation trials during the anticipatory window, which were equated in perceptual terms. Next, we studied if and how anticipation of specific stimulus categories is coded in the brain, in attention and expectation, by classifying the two anticipated categories (faces and names). Crucially, we then evaluated whether the underlying neural patterns were similar or distinct across these two preparation conditions. For this, we performed a cross-decoding analysis across both blocks, both between anticipated category decoding, and between cue and target generalization; and then as confirmatory evidence, a cross-validated representation similarity analysis (RSA). We also tested whether prepared categories engaged perceptual-like features similar to target perception, or more abstract constructs such as the encoded category (Wimber et al., 2015), by training the classifier during the anticipatory window, separately in each context, and testing during target perception. Finally, we investigated the behavioral consequences of neural anticipation by correlating decoding results to behavioral performance estimates (accuracy and reaction times).

Based on the differences between attention and expectation observed during anticipation (Peñalver et al., 2023) and target processing (e.g. Gordon et al., 2019; Zuanazzi & Noppeney, 2019) our overall hypothesis was that relevance and probabilistic preparation would lead to context-specific preactivations. At the neural level, we expected to be able to decode the different contexts (attention and expectation) in several regions. Next, we hypothesized that attention and expectation would show the preactivation of anticipated contents in visual regions, most likely in the ventral stream, but engaging at least partially separable neural patterns. These preactivations would likely show similarities to target decoding, and these similarities would be higher in the attention condition (Peñalver et al., 2023). Finally, we expected both block and category decoding to be related to behavioral performance.

4.3. Methods

Methods are reported in accordance with the COBIDAS protocol (Nichols et al., 2016).

4.3.1. Participants

Forty-six participants (mean age = 21.98, range = 18-30; 23 women, 23 men) from the University of Granada were recruited and received 20-25 euros as compensation, depending on their performance. They were all native Spanish speakers, right-handed, with normal or corrected vision. The study was approved by the Ethics Committee for Research with Humans from the University of Granada, and all participants signed informed consent prior to participation. Besides, to comply with COVID-19 safety guidelines, they wore a face mask during the whole session, including the behavioral practice outside the scanner. Six additional participants completed the task but were discarded, two due to poor behavioral performance (<80% in any of the main conditions, attention or expectation), two due to excessive head movement (either over 2.5 mm outside the voxel size or over 0.1º of rotation in 2 runs or more) and other two due to technical issues during data collection. Sample size was calculated in advance to achieve a statistical power of 80% for an estimated small effect size (Cohen's d = 0.3) and three independent variables (Block x Category x Cueing), and to match the one used in a previous experiment with a similar paradigm (Peñalver et al., 2023). Using PANGEA we obtained a minimum of 32 participants to detect the Block x Cueing interaction in reaction times and behavioral accuracy, our main behavioral prediction. Our final sample size (46 participants) provided an estimated power of 94% under the described parameters. Due to an incomplete orthogonalization of the cue-shape pairing in cases of movement in only one run, 2 participants were left out of some specific decoding analyses (n = 44).

4.3.2. Apparatus, stimuli, and procedure

Stimulus presentation and behavioral data collection were done with The Psychophysics Toolbox 3 (Brainard, 1997) on MATLAB (r2020) in a Microsoft PC. Stimuli were presented on an LCD screen (Benq, 1920x1080 resolution, 60 Hz refresh rate) over a grey background. The task, stimuli and parameters followed those used in our previous study (Peñalver et al., 2023) except that to adapt the task to the neuroimaging modality we employed longer inter-event intervals, which reduced the total trial count. We employed 160 male and female faces (50% each, extracted from The Chicago Face Database (Ma et al., 2015) plus 160 unique Spanish male and female names (50% each). Four different geometrical shapes (circle, square, teardrop and diamond with thin black outlines, unfilled) were used as cues in the main task.

The task involved a cue-target paradigm in which cues provided information about the relevance (attention) or probability (expectation) of upcoming face or word targets (Figure 4.1). Half of the blocks belonged to the attention condition and the other half to the expectation condition. To control for perceptual confounds, two cue shapes were associated with faces and two with names (counterbalanced across participants), For each participant, cue pairs changed across the experiment although their predicted category remained; the first cue for faces (e.g. a circle) appeared in half of the blocks with the first cue for names (e.g. a square) and the other half with the second cue for names (e.g. diamond). The task of participants was to indicate the sex/gender of the stimulus, responding whether or not the target belonged to the gender stated at the beginning of each block. Each block started with a screen in which they were informed about the block (attention or expectation), the target sex/gender ("Is the target male/female?"), and the two cues (one for faces and one for names). Given that attention and expectation are involved in almost any act of visual perception, we aimed at manipulating one process while keeping the other constant. In attention blocks, the cue indicated the relevant stimulus category to select (faces or names). Only if the stimulus belonged to the relevant category (50% trials, cued; e.g. is the target a female *face*?), the participant had to perform the gender discrimination task on the target. Otherwise, participants had to answer 'no' regardless of the stimulus sex/gender (non-relevant category, uncued). This manipulation of relevance, where further processing has to be applied only to selected stimuli, is similar to that employed in previous literature (e.g. Baldauf & Desimone, 2014; Saenz et al., 2002; Summerfield et al., 2006; Womelsdorf et al., 2006). Both relevant and non-relevant targets were matched in expectation, as by design they appeared with a 50% probability after each attention cue. In expectation blocks, the cue indicated the probable category of the target with a 75% likelihood (e.g. de Lange et al., 2013; Kok et al., 2017 for similar manipulations). Here, participants had to perform the gender discrimination task for both categories of stimuli, whether or not the target was cued (e.g. is the target a female *stimulus*?). This way, both the expected and unexpected targets were equally relevant. Participants were verbally instructed to use the cues in the two blocks to respond as fast as possible while avoiding mistakes.

In every trial of the main task, the sequence of events was as follows: a 500 ms cue was followed by a jittered CTI and then the target appeared for 500 ms. Both the CTI and the inter-trial duration was jittered between 2500-6000 ms intervals, spaced in 700 ms steps (average 4250). Each trial lasted on average 9.5 seconds and each run 7.6 minutes. The experiment was composed by 8 blocks of 48 trials each, or 384 trials in total. Attention and expectation blocks appeared in separate runs in a fully alternated order, and the condition of the first block was counterbalanced across participants. Cues and target stimuli were also fully counterbalanced across participants. In total, the whole experimental session lasted 60 minutes approximately, plus additional practice outside the scanner.

4.3.3. Data acquisition and preprocessing

A single session of imaging was carried out using a 3T Siemens Prisma MRI scanner, equipped with a 64-channel head coil. T1-weighted anatomical images were obtained using a rapid acquisition gradient echo (MPRAGE) sequence (TR = 2250 ms, TE = 4.18 ms, TI = 900 ms, flip angle=9°, voxel size = 1 × 1 × 1 mm). In addition, two field map images (phase and magnitude) were collected to correct for magnetic field inhomogeneities (TR = 520 ms, TE1 = 4.92 ms, TE2 = 7.38 ms, flip angle=60°, voxel size = 2.5 × 2.5 × 2.5 mm). Whole-brain functional images were acquired using an echo planar imaging (EPI) sequence (TR = 1730 ms, TE = 30 ms, flip angle=66°, slice thickness = 2.5 mm, voxel size = 2.5 × 2.5 × 2.5 mm, distance factor=0%). Slices were oriented along the AC-PC line for each participant.

The experiment consisted of 8 runs, each corresponding to a block of the behavioral task. For each run, 275 volumes were acquired, with the first 4 volumes discarded from all runs. Anatomical images were defaced to ensure anonymization. MATLAB R2020 was used for preprocessing, which involved converting the raw DICOM images from the scanner into NIfTI files with BIDS format (Gorgolewski et al., 2017). For functional images, preprocessing (with SPM12 v7771) involved the following: (1) realignment and unwarping to correct for movement artifacts (using the first scan as the reference slice) and magnetic field inhomogeneities (using previously estimated fieldmaps); (2) slice timing correction; (3) coregistration with T1 using a rigid-body transformation and normalized mutual information cost function with 4th degree B-spline interpolation; (4) registration to MNI space using forward deformation fields from segmentation with 4th degree B-spline interpolation and MNI 2mm template space; (5) smoothing using an 8mm FWHM kernel. Multivariate analyses were performed with the unsmoothed, individual subject's functional data space. Resulting images were later re-registered to the MNI space, smoothed and masked before second-level analyses.

4.3.4. Analyses

Behavioral

The main task design had three within-subject factors: block type (attention vs. expectation), cueing (cued vs. uncued) and stimulus category (faces vs. names). We calculated three-way repeated measures ANOVA for behavioral accuracy and reaction times (RTs) employing JASP (Love et al., 2019). For each participant and condition, trials with longer or shorter RTs than the average ± 2 SDs were discarded (11.5% on average).

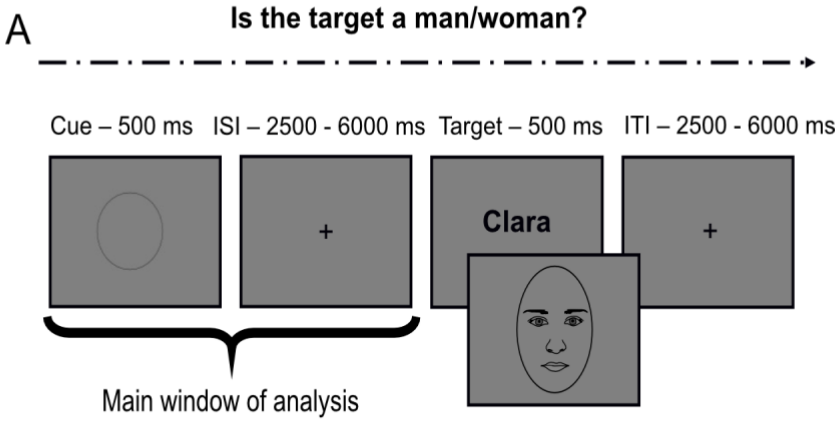


Figure 4.1. Behavioral paradigm. We employed a gender judgment task embedded in a cue-target paradigm. Participants had to respond with an index finger to indicate whether or not the target belonged to the gender indicated at the beginning of the block.

General Linear Model (GLM)

A GLM was implemented to obtain activity changes across conditions, and to obtain the betas to be used in subsequent multivariate analyses. We included cue and target regressors in the model. Cue regressors were modeled using presentation time (500 ms) plus the subsequent jitter for each trial. They were divided by Block (attention and expectation), Category prediction (Faces and Names) and Cue shape (Shape 1 and Shape 2). Importantly, although there was a total of 8 different cue regressors, these were distributed across the different runs. That is, in one run/block, cues could only be of a particular condition (e.g. Expectation), and the cues predicting a certain category had the same shape during the run (e.g. Face-shape 1 and Name-shape 2). Hence, each run included only 2 cue regressors. Target information was modeled the presentation time of the stimulus on screen (500 ms), and consisted on regressors for the conditions of Block (attention and expectation), Category (faces and names) and Cueing (cued vs. uncued). Again, although there were 8 different regressors, they belonged to particular runs, which resulted in runs having 4 target regressors each. The model also included movement information of each participant, obtained during realignment. The regressors were convolved using the canonical hemodynamic response function (HRF).

Mass Univariate

Our only contrast of interest was oriented to examine the different activations in attention and expectation during the preparatory interval. For that, we locked the events to onset of the cue, and directly contrasted both conditions. We performed this contrast for each participant, limited to grey-matter voxels with a mask, and then obtained statistical values from a second-level analysis, where they were compared using a t-test. Significance was stablished by first selecting voxels that passed a threshold of $p<0.001$, and the cluster size to a number of voxels corresponding with $p<0.05$, FWE-corrected (Eklund et al., 2016).

Decoding

Decoding was performed using beta images from the GLM with The Decoding Toolbox (v 3.999F). In all cases, we employed a two-class decoding approach with a searchlight (sphere of 4 radii, 251 voxels). We trained a classifier (Support Vector Machine, SVM) on a subset of the data, and then tested it on a different subset. Specifically, we employed leave-one-run-out cross-validation, training the classifier in all but one run, and then tested it on the remaining one. This was repeated with all runs, and the results averaged. To ensure an unbiased classification we report balanced accuracies, which accounts for over-representation of one category over the other. Group statistics applied a one-sample t-test against zero. To correct for multiple comparisons, we identified individual voxels that passed a threshold of $p < 0.001$, and then the minimum cluster size was set to the number of voxels corresponding to $p < 0.05$, FWE-corrected.

With this approach, we performed two analyses. First, we classified attention and expectation. Since there is evidence that points to the involvement of frontoparietal regions in both attention (Dodds et al., 2011; Greenberg et al., 2010) and expectation (González-García & He, 2021), we intended to examine whether information in these regions is distinguishable between both conditions even if both are similarly activated. We used all cues (identical in shape and amount across blocks) from both categories (faces and names) to train a classifier to contrast across both conditions. Since attention and expectation manipulations appeared in different runs, for this analysis we grouped them in pairs of continuous blocks, obtaining four pseudo-runs that included one of each condition, and were subjectable to leave-one-run-out classification. Importantly, this analysis does not reveal where the anticipated category is decoded differently (since all cues are used), but in what regions the general process of anticipation takes place differently depending on the relevance vs. probability manipulation. Next, we queried the regions that carried specific anticipated content by performing the classification of predicted categories (faces vs. names) separately for attention and expectation. Specifically, for each condition, we classified all the cues (two shapes) that predicted faces versus all the cues (two shapes) that predicted names. As in the previous section, the analysis was locked on cue regressors, modeled to include the entire anticipation jitter. Note that we did not compare cued vs. uncued targets, but cues that predicted (relevant or probable, depending on the block) face vs. word stimuli within the same block. Importantly, although diminished by the use of two cues per condition and participant, fully crossed across blocks, this classification approach is not completely free from perceptual confounds due to the shape of the cues. This was controlled using a cross-decoding approach (see below).

Cross-decoding

Cross-decoding was performed to assess the extent to which different conditions shared coding patterns. That is, we trained the classifier in a particular condition, and then tested it on a different one. Significant above-chance classification suggests the existence of similar patterns of brain activity in the two conditions. The crossvalidation approach was adapted to each analysis to avoid confounds, which are detailed in the following paragraphs. Searchlight and second-level analyses were applied identically as in two-class decoding.

Our primary goal was to estimate the degree to which patterns of brain activity are shared for preparation for relevant vs. probable contents. We reasoned that if anticipatory attention and expectation recruit differential patterns of brain activity, category cross-classification should not yield any significant results. We trained the classifier with cues predicting faces vs. names in one category, and then tested it on the other one. This analysis was performed without applying cross-validation. Note that this, although not identical to the original analysis, should not mean any statistical limitation, since the datasets used for training and testing are independent.

The second goal was to examine the pattern similarity between preparation and actual target perception. Thus, we trained the classifier on cue information, and then tested it on target data, separately for each condition. We applied a leave-one-run-out cross-validation approach, by training with three cue runs, and then testing on the target of the remaining one. We did this separately for the cues in attention and expectation.

Next, to examine whether cue-target parallelism is dependent on specific block context, we cross-classified between cues and targets across blocks. We did this by training the classifier in the preparatory window in one block, and then testing in the other one, again following a leave one run out approach. We performed the analysis in both directions (cue attention- target expectation; cue expectation – target attention) and then averaged the results.

Regions of Interest (ROI) extraction

To study the relationship between neural anticipation and behavior, we used specific decoding accuracy values per subject and region of interest (ROI). We considered ROIs to be the regions to return significant decoding results during the previous analyses.

We chose specific ROIs for different analyses. To study the effect of distinguishable context depending anticipatory states, we extracted the two clusters that resulted from the attention vs. expectation decoding, frontoparietal and occipital. Then, to study the consequences of better anticipated stimulus discriminability (and consequently, representation), we focused on the regions resulting from category decoding separately for attention and expectation during the preparatory window.

ROI parcellation proceeded as follows. To avoid double dipping, we performed a leave-one-subject-out procedure (LOSO, Esterman et al., 2010). For each person, we repeated the second-level analysis while leaving that participant out of the sample, so that the particular ROI was not based on their own data. Then, the resulting clusters went through the same statistical correction described above. All participants returned the same two separate clusters. Finally, we registered the two ROIs back to each participant native space using the inverse deformation fields obtained during segmentation.

To obtain accuracy values and perform correlations with behavior, we applied ROI based decoding. First, we extracted regions of interest from previous analyses. Then, classification was performed using the same rationale explained before. Only one decoding analysis was performed on each ROI, using all voxels. Again, we chose balanced accuracy to account for possible biases due to the number of observations on each condition.

Voxel selectivity ranking

We applied a voxel selectivity ranking analysis (González-García & He, 2021; Richter et al., 2018) to unveil whether neural tuning to different stimuli (instead of multivariate activity patterns) generalized from anticipation to stimulus perception.

First, for each block we obtained 8 different conditions to rank per each voxel. These images were obtained from crossing category (faces and names) with runs (4 runs per block type). Then, for each ROI, participant and voxel, we obtained the betas associated with each condition as a measure of voxel activity during the cue, and ordered them from least to most activity inducing. Next, for each voxel, we applied the same order obtained during anticipation to target perception. We reasoned that if voxel selectivity during the cue continues to the target window, the order of the eight conditions for each voxel when applied to the target should maintain a positive slope. Note that the order of the eight conditions depended on each voxel, and was not relevant by itself, but was rather a means to order target betas and study the corresponding slope. Next, we averaged all voxels, obtaining a vector of the eight ranked values, and evaluated the slope of this vector by fitting a linear regression to the ranked parameter estimates. We obtained a slope value per participant, ROI and condition. Finally, we determined whether the slope was positive (and therefore evidence of generalization from cue to target perception) by performing a right tailed one-sample *t*-test against 0, and then used False Discovery Rate (FDR) to correct for multiple comparisons. Then, if there was a positive slope in at least one of the two conditions (attention and expectation) we compared them using a two-tailed paired *t*-test, again corrected for multiple comparisons using FDR.

Brain-behavior correlation

Correlations were performed as follows. For each ROI and participant, we obtained individual accuracy results. Then, we correlated these with decoding accuracies of other ROIs' results, mean behavioral accuracies or reaction times, depending on the analyses and theoretical question. We used JASP to obtain a Pearson's r coefficient for each pairwise correlation, and used FDR to correct for multiple comparisons.

First, we studied the relationship between specific condition representations during anticipatory states and target perception. The correlation was performed between the results of the classification between attention and expectation locked to the cue in two separate clusters (Frontoparietal and Occipital), and two other measures. On the one hand, to study the stability of these sets between anticipation and target perception, we performed another classification between attention and expectation, but locked to the target. On the other, to obtain evidence that better performance is related to better context discriminability, we used each participant's global behavioral performance (acc and RTs), separately for attention and expectation.

Subsequently, we aimed to explore the relevance of anticipatory category representations. The main measure was category classification during the cue window, separately for attention and expectation, for the specific region of interest (ROI), which was correlated to two other measures: category decoding accuracy during target perception, to obtain additional evidence of the relationship between anticipatory and perceptual pattern similarity; and behavioral results (acc and RTs) to determine whether clearer representations and preparation result in improved performance. In this case, we divided the data by cueing (cued vs. uncued trials), hypothesizing that uncued trials in both contexts would show worse behavioral performance when discriminability was better.

RSA

Although preparation in attention and expectation showed different neural patterns, category anticipation could be a common process that recruits similar regions in both conditions. To examine the overlap across both manipulations regarding category anticipation, we applied a cross-validated Representational Similarity Analysis (RSA, Kriegeskorte et al., 2008) with a searchlight procedure. We designed 2 models accounting for category (faces vs. names) and cue shape (see Peñalver et al., 2023, for a similar strategy). The first one was defined as increased similarity (higher correlation) between cues predicting faces, regardless of the actual shape. The second predicted higher similarity between the identical cues regardless of what they predicted. We designed 4 by 4 matrices (one column and row per cue shape) separately for each block. Then, for each searchlight sphere, we obtained an empirical matrix separately for attention and expectation. Then, we cross-correlated each pair of matrices, obtaining one matrix per sphere. Finally, for each matrix (which accounted for the correlation between attention and expectation) and sphere, we performed a multiple regression using the two theoretical matrices as factors. This returned one *t*-value per model. The *t*-values obtained in the searchlight were entered in a second-level analysis to perform a one-sample *t*-test against zero. Correction for multiple comparisons is identical to the one used during decoding analyses.

4.4. Results

4.4.1. Behavioral results

Participants' overall performance showed high accuracies (M = 0.93, SD = 0.05). To assess behavioral effects, we used a three-way repeated measures ANOVA (Figure 4.2) on behavioral accuracy (Supplementary Table 4.1) and reaction times (Supplementary Table 4.2).

Behavioral accuracy only showed a main effect of Category ($F_{45,1} = 13.02$, $p < 0.001$, $\eta p2 = 0.22$), with less accurate responses to faces than to names (M = 0.92, SD = 0.06 vs. M= 0.94, SD = 0.07). Crucially, there was no main effect of Block ($F_{45,1} = 0.31$), indicating that attention and expectation conditions were equated in difficulty. Although the interaction between Block and Cueing was not significant ($F_{45,1} = 1.98$, $p>0.05$, see supplementary Table 4.1 for the complete result), given our hypothesis of better performance in cued that uncued trials in the expectation condition, we performed the planned comparisons and observed better accuracies for expected vs. unexpected targets ($t_{45,1}=2.75$, $p = 0.004$, Cohen's d = 0.41, M = 0.94 vs. 0.91), which was not observed in the attention condition ($t_{45,1}=0.08$, $p = 0.93$).

Only the main effect of Cueing showed a significant main effect in reaction times (RT; $F_{45,1} = 9.45$, $p = 0.004$, $\eta p2 = 0.17$), as responses were generally faster in cued trials compared to uncued trials (M = 0.701 ms, SD = 0.17 vs. M = 0.732, SD =0.18). Again, although there was no block*cueing interaction ($F_{45,1} = 1.29$, p=0.26), we studied the effect of expectations guided by our a priori hypotheses. Both conditions showed faster responses for cued than for uncued trials, but the result was only significant for expectation ($t_{45,1}=3.14$, $p = 0.014$, Cohen's d = 0.14, M = 703 vs. 728 ms) and a not for attention ($t_{45,1}=1.78$, $p = 0.3$, M = 720 vs. 731 ms).

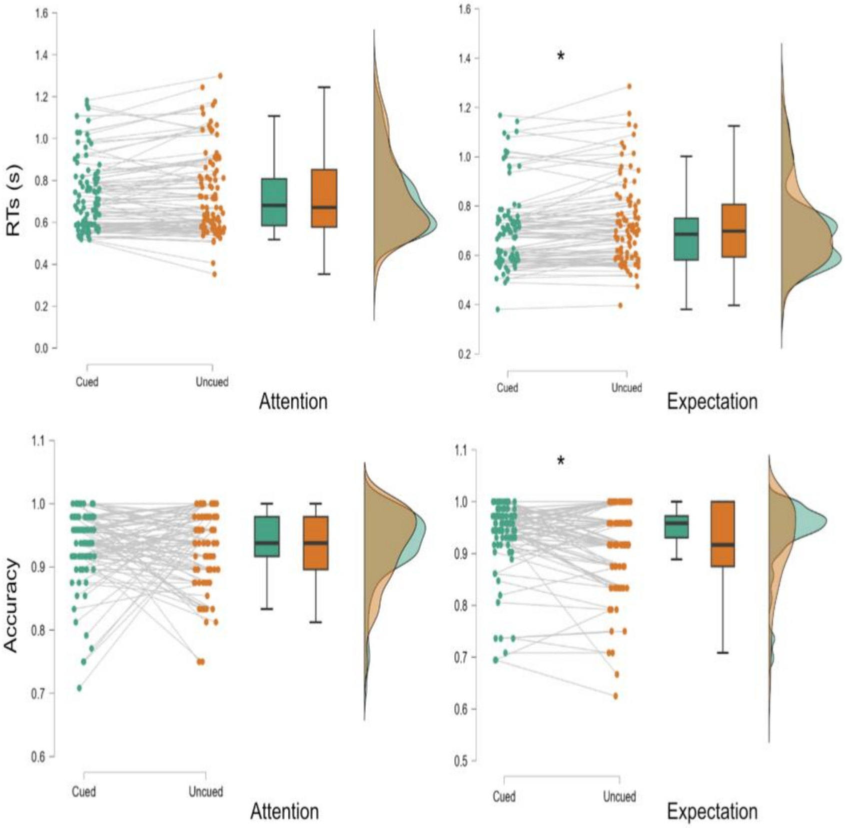


Figure 4.2. Behavioral results. (Top row) Reaction times (in seconds) in attention and expectation blocks, for cued and uncued trials. (Bottom row) Accuracy in attention and expectation blocks, for cued and uncued trials. Green sections of the plots represent cued trials, and orange sections uncued trials. Dots represent individual subjects scores per experimental condition. Grey lines connect each participant's score in the two conditions of each block. The horizontal black line inside boxes represents the median, while the limits of the box indicate the first and third quartile. Whiskers indicate the 1.5 inter quartile range for the upper and lower quartiles. Lateral plots show the kernel distribution of each data value.

4.4.2. Global differences between attention and expectation

Our first goal was to outline the regions that are involved in general anticipatory states during attention and expectation, either in global activations or in different patterns of activity. To do so, in the first place we performed a univariate contrast for attention > expectation and vice versa, always locked to the cue and for the duration of the anticipatory window (Table 4.1, Figure 4.3A). We found small univariate differences between them. Preparation in expectation induced increased activation in the posterior cingulate cortex (pCC), while V2 showed increased activations for attention.

Since previous literature has linked frontoparietal regions to anticipation in both conditions, we reasoned that even with similar activation values, the patterns of voxel-wise activations should be different if attention and expectation lead to different effects. We studied this by performing a searchlight decoding of expectation and attention locked to the cue. Our results showed significant classification in two different clusters (Table 4.1, Figure 4.3B). The first one includes several frontoparietal regions associated with the multiple demand network (MD), including the intraparietal sulcus (IS), the Supramarginal gyrus (SMG), Superior Parietal Cortex (SPC), the left Insula and the DLPFC, with the peak voxel being in the Supplementary Motor Area (SMA). The second one includes broad visual regions, including the occipital gyri (OcG), with the peak voxel being at the left lingual gyrus (LiG). Note that both blocks were perceptually identical and were also equated in speed and accuracy.

Table 4.1. Univariate contrast results

Contrast	Region	Coordinates	Cluster size (k)	Z	p	T(peak)
Cue Att.> Cue Exp.	V2	8, -80, -12	216	3.62	0.04	3.94
Cue Exp.> Cue Att.	rpCC	2, -16, 40	367	3.9	0.004	4.3
	lpCC	-4, -42, -28	210	3.84	0.044	4.22

Notes: Region labeling based on the Julich-Brain atlas (Amunts, 2020).P-values are cluster values corrected for multiple comparisons. V2 = Secondary visual area, rpCC = right posterior Cingulate Cortex, lpCC = left posterior Cingulate Cortex. Att. = attention, Exp. = expectation.

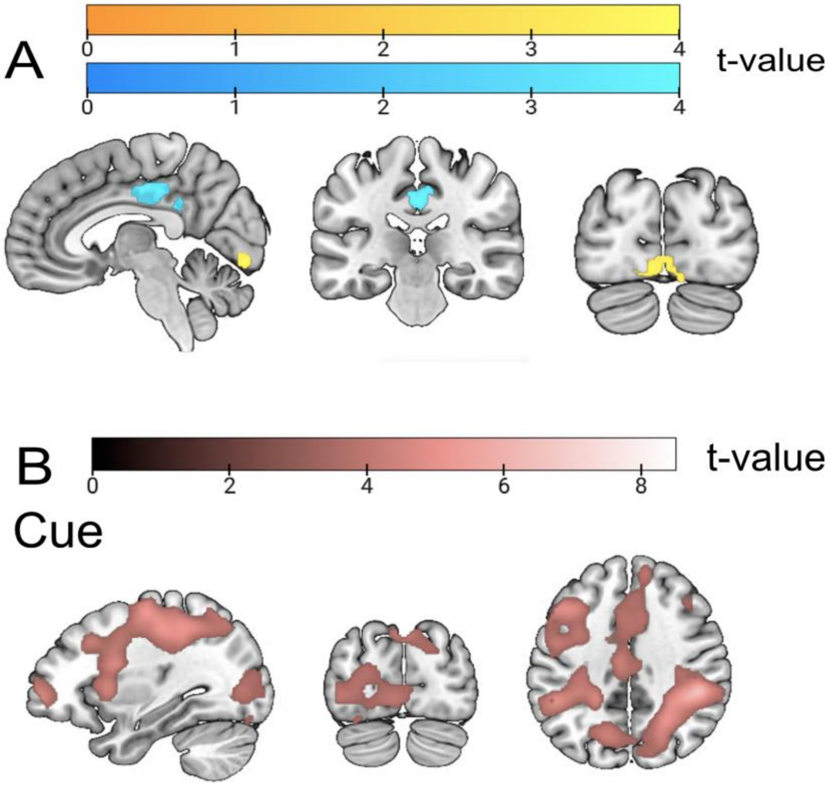


Figure 4.3. General attention vs. expectation results. (A) Contrast GLM results for cue-locked attention vs. expectation trials. Scales reflect t-values. Yellow areas highlight significant clusters for attention>expectation trials, while Blue shows results for expectation>attention. (B) attention and expectation decoding during the anticipatory interval.

4.4.3. How do block differences relate to target processing and behavioral performance?

We hypothesized that better discriminability of task related activity (block differences) during preparation should 1) extend to the target period and 2) lead to better behavioral performance.

To examine the stability of condition decoding between preparation and target perception, we performed a cross-decoding analysis, by training the classifier in the preparation interval, and testing it during target perception. However, to our surprise, there was no evidence of generalization. Nevertheless, although patterns of activity did not generalize, anticipatory condition differences could be related to the level of classification accuracy during target processing. We performed a decoding analysis locked to the cue, and then the target, in the extracted significant ROIs from the searchlight preparation analyses. Specifically, we used the occipital and the frontoparietal ROIs for each participant after applying a leave-one-subject-out procedure to avoid issues with the independence of the data (i.e., double dipping). First, we found that condition decoding during the cue was highly correlated with discriminability during target processing for both ROIs, (Occ: r = 0.378, p=0.011; FP: r = 0.44, p = 0.003), suggesting that, although subserved by different patterns of activity, the level of block discriminability is constant during the whole trial (Figure 4.4, top row). Note that these results were unlikely due to cueing effects drifting into the target window, since cross-classification was not significant.

Block decoding during the cue was also correlated with behavioral performance. Specifically, the frontoparietal ROI was positively correlated with behavioral accuracy in both attention (r = 0.396, p = 0.008) and expectation trials (r = 0.315, p = 0.037). On the other hand, although all correlations with RT had a negative sign (suggesting faster responses as decoding improved), they did not reach statistical significance (all p>0.29). The Occipital ROI showed the same pattern numerically, being positively correlated to behavioral accuracy and negatively to reaction times in both blocks, but no contrast reached significance (all p>0.11).

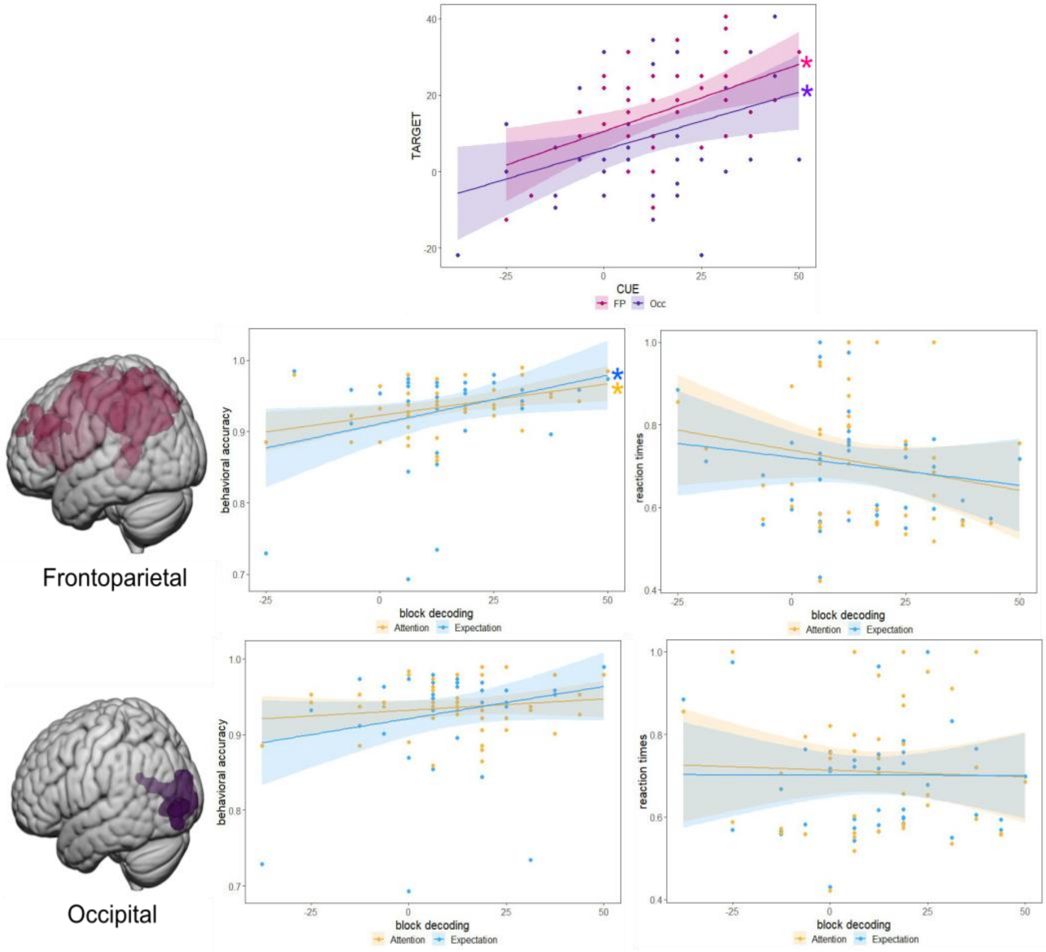


Figure 4.4. Anticipatory block differences and behavioral performance correlations. First row shows the correlation between block decoding in cue period (x axis) and the target (y axis). Second and third rows depict the correlation between block classification in the cue and behavioral results. Shaded areas show the confidence interval.

4.4.4. Differences in anticipatory category representations in attention and expectation

After studying general differences associated with the two conditions, our main interest resided in finding out the profile of visual anticipation separately for attention and expectation. Results show (see Figure 4.5A and Table 4.2) significant coding of the anticipated relevant (attention, orange) target category in two bilateral clusters in the Visual Ventral stream, with peaks at the left Inferior Temporal Gyrus (lITG) and the right ITG (rITG). In contrast, probable (expectation, blue) category anticipations showed decoding mostly restricted to the early ventral visual cortex, including the Fusiform gyrus and peaking in the Lingual gyrus.

It is important to clarify that these results are unlikely due to perceptual confounds, given that decoding was performed including two different cues for each category, which were differentially paired with the two cues of the other condition across blocks.

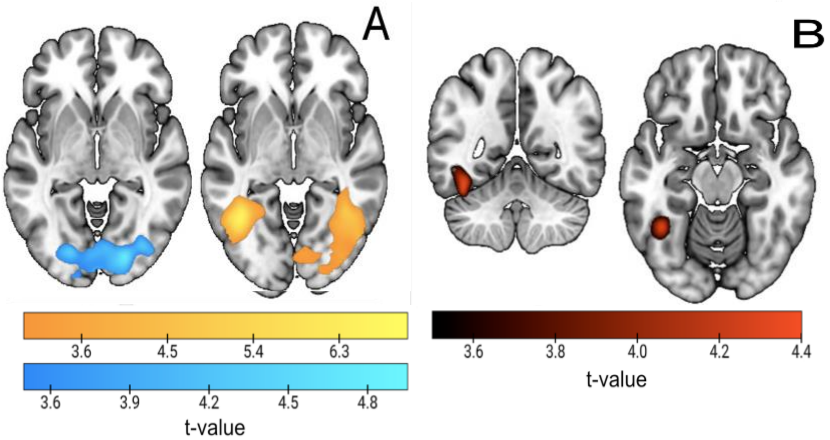


Figure 4.5. Cue-locked face vs. name MVPA. (A) Results in different blocks. Attention results are shown in orange, and expectation in blue. (B) Cross-classification results between attention and expectation.

Table 4.2. Decoding accuracy results.

Contrast	**Region**	**Coordinates**	**Cluster size (k)**	**Z**	**p**	**Decoding (peak)**
Cue attention vs. expectation	SMA	-6, -20, 50	27841	5.46	<0.001	60.34%
	LiG	-6, -76, -6	2302	3.86	<0.001	57,69%
Target attention vs. expectation	PRG	34, -18, 52	54668	6.40	<0.001	62.21%
	STC	-48, -40, 2	1595	4.93	0.002	56,44%
Cue category decoding - attention	lITG	-42, -58, -6	1564	5.73	0.001	59.91%
	rITG	42, -58, -10	1439	4.84	0.001	57.31%
Cue category decoding - expectation	LiG	-20, -84, -8	2761	4.4	<0.001	57.46%

Note: Decoding accuracy (balanced accuracy) shows the accuracy value in the peak voxel. SMA = Supplementary motor area, LiG = lingual gyrus; PRC = Precentral Gyrus, STC = Superior Temporal Cortex, ITG = inferior temporal cortex, l and r = left and right.

4.4.5. Category anticipation shows little generalization between attention and expectation

The results in the previous section show that the anticipation of categories recruits dissociable regions across attention and expectation, with overlap restricted to the right fusiform gyrus. To test whether the underlying codes are shared across relevance and probability contexts, we performed a cross-classification analysis between these conditions. One cluster (k = 756, Figure 4.5B and Table 4.3) in the left FG was significant. No other region returned significant anticipatory cross-decoding.

Importantly, even though attention and expectation showed little generalizable anticipatory decoding, within-condition analyses revealed the involvement of partially similar regions. Hence, we designed a cross-validated Representational similarity analysis (RSA) to examine the degree of overlap across both manipulations regarding category anticipation. The results are shown in Figure 4.6 and Table 4.4. A total of three clusters were found to represent category across attention and expectation. The first one was in the occipital cortex, peaking at the occipital pole (OCP), the second at the Supplementary Motor Area (SMA) and a third one at the right SMG. This result extends the findings of category decoding, and support the idea that although

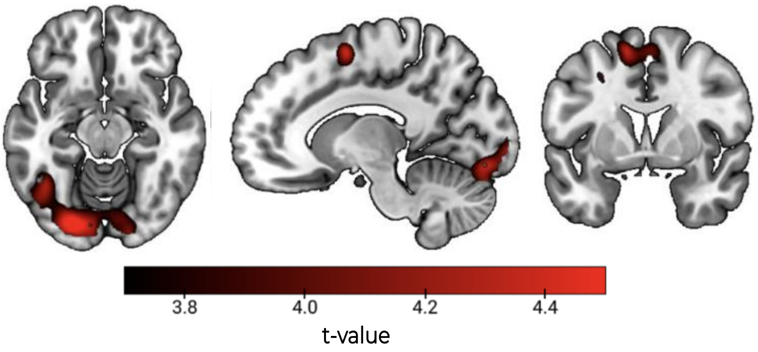


Figure 4.6. RSA Category model cross-validated across blocks. The model predicted differences across anticipated face and name stimuli. The results show the regions that were correlated across attention and expectation.

both conditions represent category in partially overlapping regions, the subserving patterns of activity are mostly unique to each context.

4.4.6. How is category anticipation related to target perception in each condition?

Results in the previous sections show that category anticipation in attention and expectation similar but not completely overlapping regions, with mostly different coding patterns. To explore whether this is due to whether anticipatory representations in both conditions differ in their level of similarity with target perception, we investigated the potential pattern overlap between preparation and the actual perception of face and name stimuli, separately for attention and expectation. We trained the classifier during cue processing and tested it during target processing (Figure 4.7 and Table 3). In attention (orange clusters in Figure 4.7), the lITG and the left Supplementary Motor Area (lSMA) showed common patterns for both processing stages. For expectation (blue clusters), results were limited to bilateral Occipital Gyri (OcG), peaking at the left Inferior Occipital Gyrus (lIOG). We performed a conjunction analysis, which only returned a cluster in the left FG, matching the coordinates of the significant regions of the cross-classification analysis (see Figure 4.4B).

However, even when most regions showed more similar representations between anticipation and target perception in only one specific condition, these differences might be explained by different levels of involvement, but not necessarily completely different representations. We investigated this by performing a cross-classification between cue and target across attention and expectation. That is, we trained with the cue in one condition, and then tested in the target of the other condition, and vice versa. We then averaged both directions. The results are shown in Figure 4.8 and illustrate the involvement of several regions of ventral visual cortex, as well as left SMA. Noteworthy, the resulting bilateral cluster included all the areas that where significant separately for attention and expectation (compare Figure 4.8 with Figure 4.7A).

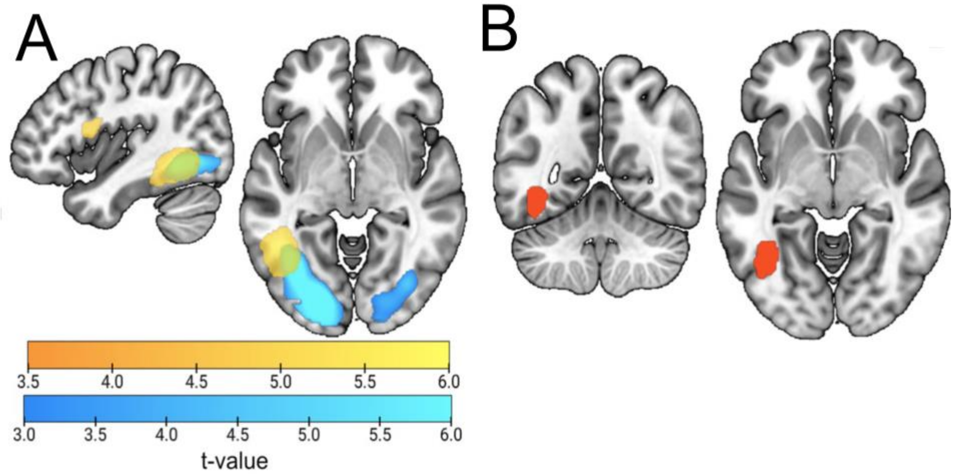


Figure 4.7. Cue-Target Cross-decoding results. (A) Attention results are shown in orange, and expectation in blue. (B) Common significant voxels across conditions (binarized conjunction).

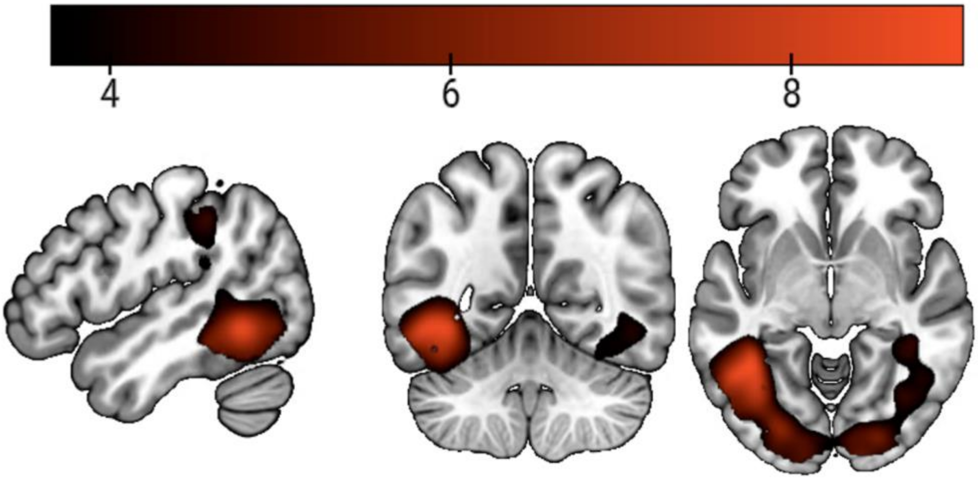


Figure 4.8. Cue–Target cross-decoding between blocks. Only one cluster was significant, but it included a big part of the regions that were significant in both attention and expectation cue-target cross-decoding.

Table 4.3. Cross-decoding results

Contrast	Region	Coordinates	Cluster size (k)	Z	p	Decoding (peak)
Cue category between blocks	lFG	-42, -50, -22	756	4.01	0.026	54.13%
Cue-Target category – Att.	lITG	-48, -52, -8	1246	5.29	0.002	57.13%
	lSMA	-52, 6, 20	594	4.17	0.03	54.17%
Cue-Target category – Exp.	lIOG	-38, -62, -2	8235	6.64	<0.001	57.12%
Cue-Target between blocks	lFG	-40, -50, -6	11993	6.71	<0.001	57.07%

IFG = Inferior Frontal Gyrus, lITG = left Inferior Temporal Gyrus, lSMA = left Supplementary Motor Area, lIOG = left inferior occipital gyrus, lFG = left Fusiform Gyrus. Att. = attention, Exp. = expectation.

Table 4.4. RSA result

Contrast	Region	Coordinates	Cluster size (k)	Z	p	T (peak)
Cue Category	LiG	-26, -96, 8	4810	4.92	<0.001	5.74
	SMA	-14, 2, 60	4447	4.18	<0.001	4.32
	rSMG	58, -46, 38	903	3.92	0.028	4.32

LiG = Lingual Gyrus, SMA = Supplementary Motor Area, rSMG= right Supramarignal Gyrus

A mechanistical explanation for the similarities found between preparatory and perceptual representations might be accounted for by neurons in both epochs of the trial being tuned to similar stimuli. We studied the stability of this neural tunning from cue to target stimuli in a univariate manner, which also allowed direct comparisons between attention and expectation. We established a voxel selectivity ranking during anticipation, and regressed it onto image preference during the target (see Methods, and Richter et al., 2018 for a similar approach). The rationale is that a positive slope of the ranked image preference during the cue, applied to target activity should act as an indicator of similar selectivity during cue and target. Conversely, a negative or flat slope would imply no univariate generalization from cue to target. We performed this analysis in the four ROIs obtained from category decoding during the cue, and cue-target cross-decoding (Figure 4.9). We separately analyzed the slope of each condition. The slope was only significantly positive for expectation, in the ventral visual ROI associated with cue-target cross-decoding for expectation, which included several areas of the VVC ($t_{43,1}=3.27$, p = 0.001), which was also significantly more positive than attention ($t_{43,1}=2.86$, p = 0.004). This result suggests that voxels within the VVC are tuned in a similar way during cue and target epochs, but only during expectation trials.

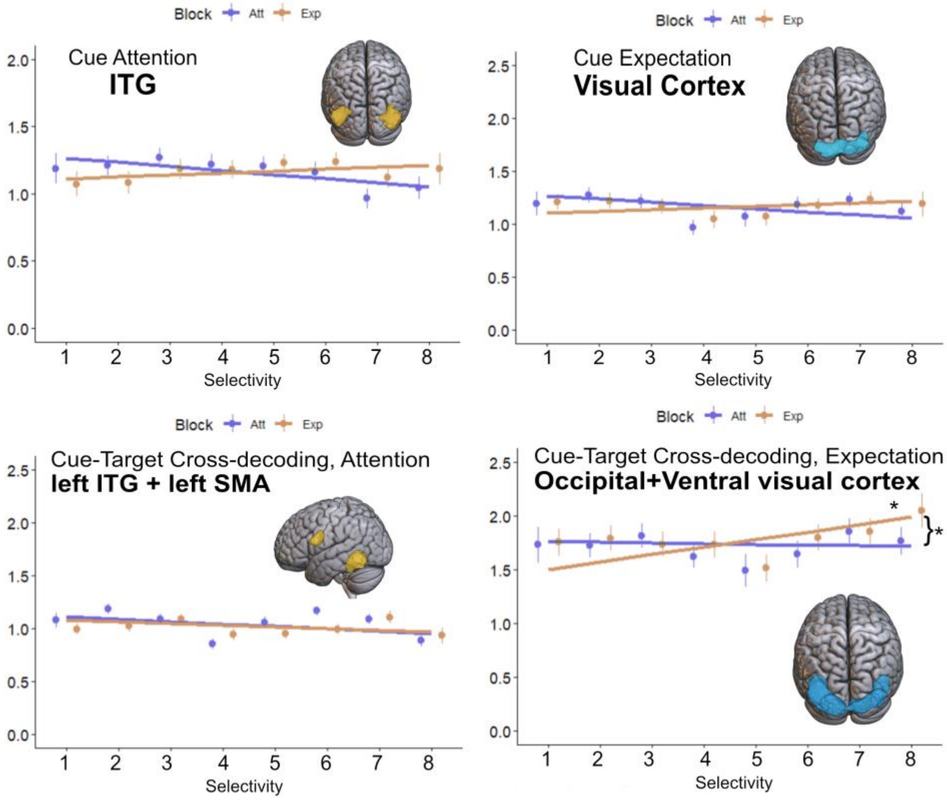


Figure 4.9. Voxel selectivity results. Each plot depicts the results for one ROI, titled as the analysis it comes from. Y axis show average beta values across all voxels and participants. X axis show selectivity preference from least (one) to most (eight) of the eight conditions during anticipation, which was then applied to target activity (see methods). Dots show mean voxel and participant beta values, and vertical lines the SD. Continuous lines indicate the slope obtained after fitting a linear regression to each condition. Asterisk indicate statistical significance.

4.4.7. Anticipatory category decoding does not correlate with target perception or behavioral results.

Finally, we sought to explore whether better anticipated category differences (classification between faces and names in the preparation interval) had an effect in 1) target category decoding and 2) behavioral performance. Overall, there was no measure that was significantly correlated to the anticipated category decoding for either attention or expectation. The complete results can be found in Supplementary Table 4.3.

4.5. Discussion

In this work we studied the specificity of top-down preparatory processes, focusing on the anticipation of relevant (selective attention) and probable (perceptual expectations) information content. We fulfilled our aim of characterizing the basis of anticipation separately for two relevant cognitive contexts, and replicated and extended the findings of previous results. Crucially, our findings show that top-down preparation is highly specific, evidenced by relevance and probability contexts leading to distinguishable brain states that extend from cue to target processing. Moreover, although anticipated stimulus categories could be decoded before target presentation in both contexts, their preparatory patterns were mostly unique for attention and expectation. Overall, these results indicate that the preactivation of information is dependent on the context, and challenge the idea that templates in selective attention and perceptual expectation are equivalent in brain functional terms.

Our task successfully manipulated relevance and probability of informational content. The task generated an overall high accuracy, indicating that cues were used in an effective manner. Importantly, expectation results showed typical cueing effects (de Lange et al., 2018; Peñalver et al., 2023; Sy et al., 2014), with improved responses for expected trials. Even though the behavioral task was adapted to match the requirements of an fMRI experiment having a reduced number of trials, it produced similar results to previous instantiations

(Peñalver et al., 2023), accounting for the robustness of expectation effects on behavior.

Our first goal was to study whether attention and expectation contexts lead to different anticipatory states. We found univariate evidence of differential anticipatory cue-generated coding of different task sets, in line with previous results (González-García et al., 2018; Hebart et al., 2018; Peñalver et al., 2023). Crucially, this could not be due to differences in perceptual factors, as blocks were fully equated on this respect, or to task difficulty differences between attention and expectation, since behavioral results were not better in any condition. Multivariate results extended these findings, showing that although activation levels where similar in attention and expectation in frontoparietal regions (Greenberg et al., 2010; Summerfield et al., 2006), the patterns that subserved each condition differed. We found two clusters, one comprising a wide range of left occipital areas, and another one in frontoparietal sites. The occipital cluster extended from earlier visual regions to areas of the left ventral cortex, while the frontoparietal part included several regions typically associated with the Multiple Demand network (MD), such as the insula, SMA, SPC and MFC (Dosenbach et al., 2007; Wen et al., 2018). This network has been related to several factors key to attention and cognitive control, including different memory load (Manoach et al., 1997) or task switching (Wager et al., 2004), although these factors were equal across our blocks. Overall differences are more likely related to the specific links between the stimulus content anticipated and the function of that content (relevance vs. probability) for the selection of the appropriate response (e.g. Woolgar et al., 2015).

Another source of useful information for understanding the anticipatory coding of relevance vs. probability is the overlap of these processes with target processing. To investigate this, we analyzed the coding stability from anticipation to perception and the behavioral consequences of these differences. For this, we performed whole-brain block decoding during target included broad frontoparietal regions. This returned larger frontoparietal regions, but, unlike during anticipation, attention and expectation could not be decoded in occipital voxels. This partial overlap could be due to similar regions being involved in both periods, but through different underlying mechanisms. To clarify this, we performed a cross-decoding analysis between cue- and target-locked activity. Contrary to our hypothesis, this did not return any significant cluster, indicating that anticipatory mechanisms that are distinctly involved in attention and expectation change throughout the trial. This, however, does not imply that anticipation is not affecting brain activity in later stages. The correlation between the average decoding in the regions during anticipation with decoding during perception was significant, despite the absence of cross-classification. Thus, it seems that the mechanisms associated with attention and expectation are deployed in similar regions, but in a different manner. Interestingly, even when patterns differ between preparatory and implementation stages, their function is related throughout the whole trial. In addition, pairwise correlations between decoding in the two clusters and behavioral performance revealed that only classification accuracies in the frontoparietal regions correlated significantly with accuracy in both attention and expectation. Altogether, results suggest that higher pattern discriminability between attention and expectation could be linked to a better representation of task demands, which would translate to improved performance (Pelzer et al., 2022).

Anticipatory representations of incoming stimuli are a widespread phenomenon found in several contexts (e.g. Hebart et al., 2018; Nobre & Serences, 2018). After learning that attention and expectation lead to distinguishable anticipatory states, our goal was to investigate whether the representational nature of relevant or probable category information about incoming stimuli was also different. We approached this inquiry by decoding anticipated stimulus category (faces vs. names) separately in each context. If the subserving mechanisms of category anticipation differs, we could expect two types of results: 1) the anticipated category is decoded in different regions depending on the context; 2) the patterns that allow to classify the anticipated category do not generalize across contexts. Regarding the first set of results, we found that it was possible to decode the anticipated category in both contexts, but the regions associated with this effect differed for the most part. Stimulus category in attention was decoded in regions of the ventral visual cortex, mostly the bilateral ITG. This region has been associated extensively with object recognition, including faces (Verhoef et al., 2015) and words (Willems et al., 2016). Importantly, there is also evidence of anticipatory processing in this region (Willems et al., 2016). The expectation manipulation, in contrast, led to category decoding in earlier regions of the ventral stream including the lingual gyrus, associated with text visual processing (Mechelli et al., 2000). Interestingly, attention and expectation differed in the anterior-to-posterior location of the anticipatory representations, with the former being in higher regions of the hierarchy, while the latter engaging earlier perceptual sites. This is similar to what was found by Kok et al. (2016), where a spatial attention manipulation had larger effects on higher order visual areas (V2, V3) while expectation was detected in V1. The reasons for this are unknown and could be partially due to the characteristics of the paradigm used, although it is suggestive of attention codes being implemented with more category-based representations, while expectation being limited to more basic perceptual coding of anticipated stimuli. Crucially, we did not decode target stimulus for cued or uncued trials separately, since this would mix reactions to matched or unmatched predictions. Additionally, by using two different cues to anticipate each category, which were fully crossed across blocks for each participant, makes it highly unlikely that the significant decoding was due to the perceptual features of the cues.

In addition, to query the similarity of the patterns, we implemented a cross-classification analysis between the attention and expectation. There was scarce evidence of cross-decoding, which was limited to a small cluster in the left Fusiform Gyrus. This adds support to the idea that the anticipation of specific relevant vs. probable stimulus categories is mostly based upon different mechanisms, replicating and extending previous results (Peñalver et al., 2023). Finally, we run an RSA (Kriegeskorte et al., 2008) to find the regions that coded for category across conditions, regardless of the specific patterns that may be driving the dissociation found in the previous decoding. This analysis showed a combination of the regions that had appeared during decoding analyses performed separately in attention and expectation, including LiG and ITG together with premotor and inferior parietal areas. Moreover, other regions such as SMA also showed evidence of anticipatory category coding, although they were not significant in the decoding of either condition. This adds support to the notion that some frontoparietal regions also represent anticipated information, although more trials of each condition might be needed to find evidence separately in each condition. Altogether, this result provides further evidence in favor of anticipatory representations sharing representational characteristics in attention and expectation, while, at the same time, being subserved by distinct patterns of brain activity.

A potential source for differences between attention and expectation could be that they engaged different levels of similarity between anticipated and perceived target information, similarly to what has been suggested for visual imagery (e.g. Cichy et al., 2012). Both cross-classification between cue and target was evident in both conditions, however these differences appeared in mostly different areas. Attention showed cross-decoding on the left ITG, plus in the IFG. On the other hand, this analysis engaged more anterior occipital and temporal ventral regions for expectation. Surprisingly, significant regions did not fully match the ones found during anticipated category decoding, which suggests that anticipatory coding was taking place also in these areas but did not reach statistical significance. Interestingly, there was a small region in the fusiform gyrus that appeared in both attention and expectation (see Figure 4.7B), which matches the only significant cluster obtained during cue cross-decoding, and also part of the RSA analyses. Altogether, this suggest that at least the left FG might uniformly act in different preparation contexts.

Importantly, our results suggested that anticipatory templates in both conditions share some similarity with stimulus perception, regardless of the condition. This was evidenced by a cross-classification analysis in which we trained the classifier with cues with one condition, and then tested it in targets of the other. The result closely matched regions that combined the ones that appeared when studying each condition separately. Since stimulus perception arguably leads to strong perceptual patterns, they could generalize to different forms of anticipatory processing. In this regard, Cichy et al. (2012) proposed that perception involves a larger set of effects than, in their study, imagery. In our case, the processes involved in perception in either attention and expectation might actually include activity similar to anticipation decoding in both conditions, while anticipation in each condition might be more specific.

A voxel selectivity ranking analysis provided further evidence regarding the degree of generalization of neural patterns from cue to target stimuli. We did this in the regions that so far were found to be related to category anticipation: cue decoding and cue-target cross-decoding for attention and expectation. Out of the four regions, we only found evidence of generalization in the cue-target region for expectation, in broad ventral stream sites. There, only expectation showed evidence of generalization, and it was also significantly larger than for attention. Altogether, this set of results highlights that preparation for different categories occurs in a representational manner that is akin to stimulus perception. Although the phenomenon is increased in different regions depending on the context, it appears to be at least partially common in the different contexts. Crucially, univariate evidence only supported this generalization for expectations, suggesting that cue-target similarities in probability contexts might be due to voxel tuning to the anticipated category, while selection might achieve these representations by different mechanisms.

As a last step, we investigated whether category anticipation could be related to the efficiency of behavior. In contexts of expectations, classification accuracies have been previously reported to correlate with behavioral performance (Kok et al., 2017; Van Ede et al., 2012). However, and contrary to our hypothesis, we did not find such relationship in either attention or expectation. The reasons for this are unclear. Notably, in a previous analysis, we found that global condition discrimination did correlate with behavioral accuracy. Hence, whilst it could be the case that general condition preparation has a more direct connection with motor responses, the specific mechanisms that allow for content maintenance in visual regions could be related to other processes that are less influential on behavioral performance. Moreover, since RSA showed evidence of category anticipation for both regions in SMA that was not included in this correlation analysis, future studies that specifically target object representations in frontoparietal regions might find a closer relationship with behavior.

Our results add to the evidence that attention and expectation engage neural anticipation, although how this happens is still not fully understood. The current evidence could align with predictive processing accounts of the anticipatory effects of both contexts (Auksztulewicz & Friston, 2016; Feldman & Friston, 2010; Kok et al., 2017). Attention consistently showed anticipatory coding in regions of the ventral visual cortex that are related to precise stimulus perception, while expectation was often related to earlier visual areas. In this scenario, selection would act as a closer representation of the relevant stimulus categories, perhaps more closely related to brain imagery (Christophel et al., 2015; Lawrence et al., 2019). This could happen by preparatory increases of gain in neurons tuned to relevant categories in order to enhance information sensory weights when the target appear (Feldman & Friston, 2010). On the other hand, expectation would induce an excitability increases in less complex regions tuned to basic stimulus features, allowing for a more flexible representation and induction of prediction errors to reduce noise during visual processing and in case of unfulfilled expectations. Importantly, this is further supported by the results showing how voxel tunning remained constant only in expectation, adding to the concept of neural excitability and responsiveness to probable stimuli.

Although overall our results account for differences found in related processes such as attention and expectation, they should be complemented by future research. Importantly, our scope was limited to the anticipatory interval and its similarities with perception. We chose to study the similarities between anticipation and stimulus perception by using targets from the same blocks as the cues. This could lead to partial overlaps between cue and target perception due to the slow changes in BOLD signal, that may partially account for the similarities between the two parts of the trial. Although these results replicate and complement the ones previously found using EEG, which is temporally precise, an unbiased spatiotemporal profile might be obtained by applying multimodal fusion methods (Cichy & Oliva, 2020), which allow to combine the information supported by both techniques. Moreover, to make an unbiased comparison, the best strategy is to obtain data from an independent localizer block without task demands or response noise that would allow to implement a canonical template tracking procedure. This would give more clear evidence of whether anticipatory information was akin to target perception (González-García et al., 2018; Palenciano et al., 2023; Wimber et al., 2015). However, it is important to note that, when cross-classifying cues and targets from different conditions, classification was still possible. This makes it an unlikely explanation that mechanisms involved in preparation-perception similarity are completely biased by block effects. Finally, it is still debated whether attention varies regarding the probability of incoming stimuli (Alink & Blank, 2021; Gordon et al., 2019), which makes it difficult to completely eliminate a component of attention in the expectation condition. However, our manipulation was not designed to study that attentional component, and hence our results speak regarding how they differ, but not how they influence each other. Previous research studying how attention influences expectation effects has focused on target processing, and has yielded mixed effects (Kok, Rahnev, Jehee, Lau, & de Lange, 2012; Richter & de Lange, 2019). Future studies that orthogonally manipulate both conditions might study anticipatory processing to respond to whether probable stimuli increase or decrease attention.

Altogether, with a paradigm that manipulated the anticipation of relevant or probable contents in tasks equated perceptually and in difficulty levels, our neuroimaging results show the preactivation of specific patterns is mostly different for attention and expectation. In both conditions cues pre-activated patterns similar to those reinstated during subsequent target perception. However, this overlap happened in mostly different regions, and only in expectation it was possibly due to specific neural tuning to probable stimuli. Our findings suggest that selection acts through complex stimulus representations, while expectation increases excitability in earlier, more basic perceptual regions. Crucially, we have shown that preparation is a complex phenomenon, where different cognitive processes share representational structures, although implementing advantageous effects via different mechanisms that are mostly specific to certain contexts, complemented by activity constant through different forms of preparation. Overall, our results replicate and extend previous findings, thus stressing the specificity of anticipatory processing depending on the informative role it plays for further target processing.

4.6. Supplementary Materials

Supplementary Table 4.1. Behavioral accuracy results.

Conditions	**F**	**p**	η^2_p
Block	0.311	0.580	0.007
Cueing	2.231	0.142	0.047
Category	13.024	< .001	0.224
Block ✻ Cueing	1.977	0.167	0.042
Block ✻ Category	24.112	< .001	0.349
Cueing ✻ Category	6.766	0.013	0.131
Block ✻ Cueing ✻ Category	6.717	0.013	0.130

Supplementary Table 4.2. Reaction Time results.

Conditions	**F**	**p**	η^2_p
Block	1.475	0.231	0.032
Cueing	9.451	0.004	0.174
Category	1.679	0.202	0.036
Block ✻ Cueing	1.291	0.262	0.028
Block ✻ Category	1.417	0.240	0.031
Cueing ✻ Category	0.679	0.414	0.015
Block ✻ Cueing ✻ Category	5.384	0.025	0.107

Supplementary Table 4.3. Cue category decoding correlations, separately for attention and expectation.

Attention		
	Pearson's r	**p**
Target cued	-0.163	0.290
Target uncued	-0.089	0.566
Acc. cued	0.160	0.299
Acc. uncued	0.009	0.955
RT cued	-0.287	0.058
RT uncued	-0.242	0.114

Expectation		
	Pearson's r	**p**
Target cued	0.013	0.935
Target uncued	2.896×10^{-4}	0.999
Acc cued	-0.122	0.430
Acc uncued	-0.271	0.075
RT cued	-0.020	0.897
RT uncued	-0.061	0.693

Note: Table shows the correlation values between the classification accuracies obtained during category (predicted faces vs. names) decoding locked on cue, separately for attention and expectation. The ROIs are the ones that were significant in the cue category whole brain decoding analysis. Attention ROI = ITG, expectation ROI = Early visual cortex. Target cued and target uncued refer to the category decoding result locked to the target, in the same ROIs as the cue depending on the condition. Acc and RT refer to behavioral results, separately for condition and cueing.

Chapter V
Study III

The content of this chapter is in preparation as Peñalver, J.M.G.; González-García, C.; Ruz, M. Different prior induced category representations in expectation and attention. An EEG and fMRI study.

5.1. Summary

Top-down preparation is a key tool that allows efficient behavior in a complex environment. Indeed, it is generally believed that prior-induced anticipation has an effect on target processing. The effects are usually associated to enhanced activations for attended information (vs ignored), and decreased activations for probable stimuli (vs unlike). However, the mechanisms behind these phenomena are not well understood. Specifically, it is debated whether probable stimuli dampen or sharpen sensory processing. Moreover, it is not clear whether improbable stimuli automatically receive more attention, making attention and expectation tightly connected processes. Here, we re-analyzed two previous experiments (Peñalver et al. 2023 and Peñalver et al. in prep) which independently manipulated selective attention and perceptual expectations in a similar cue target paradigm, using EEG and fMRI respectively. In different blocks, participants were cued to the relevance or the probability of face and name stimuli. Using MVPA to decode target categories (faces vs. names), we sought to add evidence to the debate on whether attended stimuli receive sharpened processing (indicated by larger decoding accuracies) and if probable targets elicit dampened (poorer classification) or sharpened representations. In the EEG experiment, we aimed to specify the temporal profile of these effects, both their onset and the temporal stability. In fMRI, we studied how frontoparietal and visual regions are involved in dampening/sharpening effects. Both experiments showed a pattern that is consistent with dampened representations for probable stimuli, and unconclusive evidence supporting a sharpening of representations for attended stimuli. Moreover, we showed that these effects have early onsets, that are temporally stable and are extended after target perception, and are implemented mostly in ventral visual regions.

5.2. Introduction

In a complex and ever-changing environment, the ability to anticipate incoming stimuli is crucial for flexible and well-adjusted behavior. Theoretical frameworks of proactive cognition view preparation as an endogenous brain function that occurs before stimulus presentation, improves behavior (Stein & Peelen, 2015), and induces neural representations of incoming information (e.g. González-García et al., 2018; Kok et al., 2017). While it is widely accepted that preparation is a beneficial process, there is limited understanding of whether it is a unified or heterogeneous phenomenon. Research has found differences in preparation for different behavioral demands, such as attentional and motor preparation (Brunia, 1999); different task sets (Hebart et al., 2018); or levels of contextual demands (Kim et al., 2020). Crucially, recent research showed how the anticipation of specific stimulus categories differed in contexts of selective attention and perceptual expectations (Peñalver et al., 2023).

However, how does preparation in different contexts influence target processing? Research has widely studied such effects of anticipated probability and relevance contexts. A common finding in literature is that perception of probable stimuli decreases neural activity (or, conversely, unexpected information increases activity), a phenomenon labeled “expectation suppression”. Electroencephalography (EEG) correlates such as the mismatch negativity (Garrido et al., 2009; Pazo-Alvarez et al., 2003) are thought to reflect a reduced processing demand by probable information compared to unexpected stimuli. Similar effects where novel stimuli generate larger EEG amplitudes than expected ones have also been extensively found (e.g. Manahova et al., 2018). Likewise, oscillatory activity appears to reflect converging findings. Such is the case for midfrontal Theta (4-8 Hz) power, associated with control (Cavanagh & Frank, 2014), which increases in contexts of target unexpectedness (Cavanagh et al., 2012). In addition, Gamma (>30 Hz) power has been shown to decrease with stimulus repetition in monkey V4 cortex (Brunet et al., 2014). Similarly, BOLD signal for expected information is reduced in regions such as V1 and lateral occipital cortex (e.g. Kok et al., 2012; Walsh & McGovern, 2018; see Feuerriegel et al., 2021 for a detailed review). These findings are usually explained under the predictive processing framework (Kok & de Lange, 2015), reflecting efficient priors which in turn reduce prediction errors.

Importantly, the nature of the neural mechanisms that generate expectation suppression is still debated. There are two main and, in principle, mutually exclusive explanations (de Lange et al., 2014; Press et al., 2020). The first one proposes a suppression of the neural responses tuned to predicted information, and therefore, redundant (Murray et al., 2004; Friston, 2005). This results in a "dampening" of neural populations tuned to predicted features, which results in a rapid updating of knowledge in case something unexpected occurs (Press et al., 2020). The second explanation proposes a decrease of activity of neural populations tuned away from expected information, inducing a "sharpening" of the representation from selective neural populations. Since these accounts cannot be immediately compared with mean population changes, multivariate analyses have been used to study such patterns of brain activity (de Lange et al., 2018). The rationale is that if expectation dampens selective neurons, the fidelity of representations in these regions should be reduced, and decoding accuracies should in turn decrease. Conversely, if expected information sharpens neural activity, the fidelity of representations would improve and thus classification accuracies should increase (Kok, Jehee, et al., 2012).

Previous research has reported supporting evidence for both the sharpening and dampening accounts (González-García & He, 2021; Han et al., 2019; Kok, Jehee, et al., 2012; Richter et al., 2018; Walsh & McGovern, 2018; Yon et al., 2018, 2023). Importantly, expectation suppression effects, although robust, have been found more consistently in tasks that utilize a small number of stimuli and generate strong associations through frequent repetitions, such as statistical learning (Richter et al., 2018). However, cueing paradigms have yielded mixed effects (Feuerriegel et al., 2021). Nonetheless, since dampening and sharpening effects refer to changes in specific neural populations, voxel-wise modulations could be studied in the absence of overall activity changes (González-García & He, 2021).

In contexts of attentional selection, EEG research has shown larger voltage amplitudes for attended vs. ignored information, for example colors (Störmer & Alvarez, 2014), positions (He et al., 2004), or words (Perrone-Bertolotti et al., 2020). Moreover, different oscillatory correlates have been associated with selection. Fronto-medial Theta power increases in contexts where cognitive control is needed, such as distractor inhibition (Cavanagh & Frank, 2014). In addition, Alpha (8-12 Hz) power and synchrony have been shown to decrease for selected stimuli (e.g. Bagherzadeh et al., 2020; Banerjee et al., 2011; Feng et al., 2017). Also, Gamma power (Gruber et al., 1999) and synchrony (Doesburg et al., 2008) are higher for task relevant vs. task irrelevant stimuli. Relatedly, studies using fMRI have often found increases in cortical activity for cued stimuli in occipitotemporal regions in non-human (Bichot et al., 2005; Burrows et al., 2014; Chelazzi et al., 1998; Cohen & Maunsell, 2011) and human (Kastner et al., 1999; M. V. Peelen & Kastner, 2011; Serences et al., 2004) primates. Altogether, these findings support classic theoretical frameworks such as the biased competition theory (Beck & Kastner, 2005; Desimone & Duncan, 1995), which usually is tested with tasks that simultaneously present several competing stimuli, and therefore may involve suppression to highlight the attended stimulus. This model contends that all stimuli in the visual field compete for representation, which is evidenced by the finding that neural responses to individual stimuli are larger than responses in the presence of several images (e.g. Kastner et al., 1998). This competition can be biased by selection, which increases cortical sensitivity to relevant stimuli, while suppressing activity of neurons tuned to ignored stimuli. Crucially, this effect is modulated by top-down attention.

The notion of suppression for distracting or irrelevant information resonates with the predictions made by sharpening models of expectation: sensory representations are increased for attended information, since neurons tuned to unattended information are suppressed. This is supported by several findings. For example, Martinez-Trujillo & Treue, (2004) observed that the response of monkey MT neurons was increased if an attended feature was close to the receptive field of the neuron, and decreased if it was unattended. Moreover, evidence of this effect has also resulted from multivariate analyses, which showed better decoding for attended vs. unattended stimuli (Goddard et al., 2022; Vaziri-Pashkam & Xu, 2017; Woolgar et al., 2015).

In the current study we investigated whether selection and probability anticipations provide results consistent with dampening vs. sharpening of neural activity to target stimuli. We analyzed data from two neuroimaging experiments designed to contrast face and word anticipation in these two conditions. The task employed a cue-target paradigm in which, in different blocks, cues gave information about relevant or probable stimuli. In the first experiment, we recorded EEG data and included an independent localizer of target stimuli. In the second, the task was adapter to collect fMRI data.

In both experiments we followed a similar analysis rationale. First, we estimated univariate effects to cued targets matching expectation suppression and attention enhancement. In the EEG experiment we contrasted the amplitude of the Event Related Potentials (ERP) of cued and uncued stimuli, and also, we performed a time-frequency decomposition. In the fMRI experiment we studied the regions that have previously been reported to show effects of expectation suppression or attention enhancement, and are related to either sharpened or dampened sensory representations; namely, visual regions such as V1, visual ventral cortex (VVC), fusiform gyrus (FG) and lateral occipital cortex (LOC); and multiple demand (MD) regions like frontoparietal cortex in the inferior frontal sulcus (IFS), anterior insular/frontal operculum (AI/FO), anterior cingulate cortex/pre-supplementary motor area (ACC/pre-SMA), and the intraparietal sulcus (IPS) (Duncan, 2010; González-García & He, 2021; Richter et al., 2018; Störmer & Alvarez, 2014; Woolgar et al., 2015). Next, in an exploratory analysis we performed a whole-brain contrast to find other regions that might be involved in dampening or sharpening effects.

To adjudicate between dampening vs. sharpening for attention and expectation, we applied multivariate pattern analyses (MVPA) to decode the target-locked categories (faces vs. names) separately for expectation and attention conditions, in both EEG and fMRI. If expected stimuli induce a dampening effect, decoding should be worse for cued than uncued trials (Blank & Davis, 2016; Kumar et al., 2017; Walsh & McGovern, 2018). However, if neural responses tuned to specific categories are sharpened, the multivariate representations should be enhanced (Bell et al., 2016; Kok, Jehee, et al., 2012; Yon et al., 2018). In attention, we expected to find evidence of sharpened neural representations, as it has been shown using a similar methodology (Goddard et al., 2022; Woolgar et al., 2015). Crucially, in the EEG dataset, we performed temporal generalization analyses (Grootswagers, Wardle, et al., 2017; King & Dehaene, 2014) to study whether dampening/sharpening markers are stable through time or reflect specific modulations of activity at different time-points after target processing. Relatedly, we also had hypotheses regarding the relationship between dampening and sharpening effects with preparatory activity, based upon two sets of findings. On the one hand, a dampening effect is supposed to attenuate sensory processing of probable or unattended perceptual representations (Press et al., 2020). On the other, preparatory brain activity has been shown to be characterized by displaying attenuated versions of target processing (Dijkstra et al., 2022; Koenig-Robert & Pearson, 2021). We reasoned that the attenuation associated with dampening would be similar to the reduced effects that arise during anticipation compared to target perception. Hence, dampened representations should show more evidence of cross-classification between cue-locked and target locked activity. Thus, in the EEG experiment we studied the relationship between these effects and pre-stimulus representations, for which we applied temporal generalization in extended epochs that included both cue and target data.

In the fMRI experiment we also performed classifications between different targets, compared decoding accuracies across the ROIs described above, and then performed an exploratory whole brain analysis to find potentially relevant evidence in different regions.

5.3. Experiment 1

5.3.1. Methods

The complete sample, task and data acquisition steps are described in Chapter 3 of the present thesis (and also in Peñalver et al., 2023).

Acquisition and Preprocessing

Acquisition and preprocessing steps are described at length in the Experiment I of the present thesis. For the analyses presented in this chapter, some changes were implemented during preprocessing. Specifically, instead of using different epochs for cue and target stimuli, we split all the EEG data into 7-second epochs that covered from cue to target stimuli. After automatic trial rejection with the new epochs (see Chapter 3 for the detailed procedure), an average of 8% trials were rejected per participant (range 1.8%-19%). All other details remained the same.

Analyses

Whole brain ERP analyses

To find evidence during target processing of cueing effects either of expectation suppression or attention enhancement, we first contrasted the event related amplitudes associated with cued and uncued targets. ERPs were analyzed at each electrode in discrete time bins using mass univariate dependent samples *t*-test, implemented in Fieldtrip (Oostenveld et al., 2011). This measure provides information regarding possible widespread effects that are not necessarily linked to specific channels. Statistical comparisons were performed between cued and uncued targets, separately for attention and expectation. Importantly, we analyzed the two categories (faces and names) independently, since they are known to generate different EEG correlates (Molinaro et al., 2013; Rousselet et al., 2004) and behavioral consequences (Peñalver et al., 2023). Time-wise, we averaged voltage values of discrete windows in the following manner: P1, 50-150 ms (Novitskiy et al., 2011); N1, 150-190 ms; P2, 190-250 ms (Tenssay & Wang, 2019), and N2, 250-450 ms (Pazo-Alvarez et al., 2003). Furthermore, to reduce computational costs and to increase power, all samples within each ERP window were averaged. Moreover, average referencing can result in biases in reporting whole-brain ERP effects, since dipole formation might induce a significant effect in a part of the scalp and opposite effects in the opposite part, in addition to a bias to artificially reduced effects in central electrodes (Junghöfer et al., 1999). Hence, we only queried electrodes that have previously been linked to the specific ERPs selected for each window. For visual potentials we kept all occipital, parietooccipital and temporoparietal electrodes from the 64-channel net. To explore effects of cueing outside of visual regions, we analyzed the P3 potential, given its involvement in conflict in selective attention and unmatched expectations (Feuerriegel, Vogels, et al., 2021; Mueller et al., 2008). We focused on central and frontal electrodes, and analyzed the window between 300 and 650 ms (Luck & Kappenman, 2011).

For each ERP window, we did as follow. First, for each participant, we averaged all the trials from each condition. Then, for each electrode we calculated the t statistic between the two conditions of interest. All electrodes with $p<0.05$ values were clustered with neighboring electrodes with akin results. Neighbors were defined using the *ft_selectneighbors* function from Fieldtrip, based on each electrode's coordinates. The cluster statistic reported is the sum of all the t-values of the electrodes included in a cluster (Δt), which is used to compute statistics. Correction for multiple comparisons was performed using a cluster-based Monte Carlo permutation test. We repeated the analyses

10000 times by extracting data from random participants and conditions per contrast. This yielded a random distribution of cluster sizes. Clusters from the original sample that were outside the 95% of the distribution were deemed significant.

Time-frequency analysis

To explore univariate effects not captured by ERPs, we analyzed time-frequency data. We performed frequency extraction separately for each participant, electrode and condition. The data were convolved with complex Morlet wavelets (logarithmically spaced in 48 frequencies from 2 to 50 Hz) and the wavelet's width was defined for each frequency by logarithmically spacing from 4 to 10 in the same number of steps. The power spectrum was estimated with a fast Fourier transform, and then multiplied by the power spectrum of the frequency specific Morlet wavelet. Then, with an inverse Fast Fourier Transform we obtained the time-resolved analytic signal. The squared absolute value of this signal returned the power values for each channel and trial. We performed the decomposition for the whole trial interval, in windows of 50 ms, to avoid limitations in slower frequencies. Statistical comparisons were performed only from -0.1 to 0.65 seconds locked to the target, as effects of suppression and enhancement are usually found in that window (Henson et al., 2003; Summerfield et al., 2011).

We again applied cluster correction for multiple comparisons. This time we kept individual time points and averaged power results across whole frequency bands (theta: 4-7 Hz, alpha: 8-12 Hz, gamma: 31-49 Hz), since our hypotheses were whole band-related and we did not have hypotheses regarding the onset of the effects. The procedure was akin to the one described in the previous section, only this time adjacent time points were also considered for clustering, along with spatial neighbors.

Decoding

We studied whether cueing improved or impaired stimulus category coding by performing a classification analysis on target data. We used all 64 electrodes and voltage as decoding features, applied a Linear Discriminant Analysis algorithm for classification to obtain the AUC. Statistical significance was estimated by a cluster-based permutation analysis (please see Chapter 3 for details).

We performed two sets of decoding analyses. In the first one, we classified face and name targets, separately for cued and uncued trials, in attention and expectation, thus creating four decoding conditions. We used windows from -100 to 1500 ms locked on the target. In the second analysis, we used whole-trial windows, ranging from cue to target processing, from -100 to 4150 ms. To avoid perceptual confounds due to the identity of visual cues, we applied the same cross-classification scheme described in Chapter 3. In short, we trained with trials starting with one pair of cues (one predicting faces and the other names), and then tested on different trials with a different pair of cues. In the two analysis we also applied a temporal generalization approach (King et al., 2014), by training and testing in every time-point. Note that the temporal generalization matrices (TGMs) resulting from the last analysis contain four sections: Only cue (train and test in the cue interval, from 0 to 1550 ms); only target (train and test in the target, from 1550 to 4150 ms); train cue - test target, from 0 to 1550 ms in the x axis, and from 1550 to 4150 ms in the y axis; and train target - test cue, with the reversed window (see Figure 5.5A). All analyses were performed using MVPAlab (López-García et al., 2021) running on MATLAB. Statistical significance was obtained by applying cluster-based permutation analyses as described in Chapter 3.

Statistical comparisons across decoding results

Statistical comparisons across decoding results were performed using cluster-based permutations, with a slightly different rationale than the ones described in Chapter 3. We used time-resolved ANOVAs and *t*-test analyses to compare cued vs. uncued decoding in attention and expectation. For each time-point we used custom code and built in MATLAB functions to perform the analyses. We first performed a 2 by 2 ANOVA (block*cueing) on every time point. This resulted on a vector with three F-values (two main factors and the interaction) per time point. Then, we generated permutations by randomly multiplying one of the conditions by -1 in each subject. We did this 5000 times. Next, we extracted cluster information in the true data by looking for sets of temporally adjacent points with $p<0.05$. For each cluster, the sum of F values was stored as the size of the cluster. We repeated this with the 5000 permutations, obtaining a distribution of cluster sizes, and used the 95th percentile to mark the minimum cluster size considered significant. Then, to perform post-hoc comparisons, we used a *t*-test between cued and uncued trials and followed the same approach (see Dalski et al., 2022; Dijkstra et al., 2018 for similar approaches). For temporal generalization analyses we did the same, but instead of using only a vector of results, we used the complete temporal generalization matrices, obtaining 3 F-values or one *t*-test result per each timepoint-by-timepoint comparison.

5.3.2. Results

5.3.2.1. Univariate results

The analyses performed to dissociate the univariate effects of cueing during target perception for attention and expectation revealed several amplitude differences between cued and uncued attention targets. Overall, uncued targets induced larger voltage values in anterior sites, while cued targets induced larger values in medial and frontal electrodes (Figure 5.1, Supplementary Table 5.1). Specifically, for names there was a cluster of larger amplitudes for uncued targets in the P1 window and the P3 left central sites, and similar results for both names and faces in the P2 and the N2 window. On the other hand, the P3 showed higher amplitudes for selected targets for both names and faces. Surprisingly, there were no significant univariate differences between expected and unexpected targets (all $p>0.05$).

Next, we contrasted whether and how cueing differently influenced oscillatory activity in contexts of selection and probability. Results are depicted in Figure

5.2 and Supplementary Table 5.2. In attention we found power differences in several bands. In Theta there were large differences between selected and non-selected stimuli, with uncued targets leading to a cluster of increased Theta that started on midfrontal sites at 50ms after target onset, and extended over the whole scalp, from 200 ms to the end of the window. The alpha band (8-12) was significant for uncued targets in a cluster that started on frontal sites at 140 ms, moved to central electrodes at 200 ms, and finally stayed on parietal and occipital channels from 450 to 650 ms, but only for faces. Names returned similar results, but clusters were more temporally separated and some parietal and occipital electrodes survived multiple comparison correction between 400 and 650. Finally, we found significant differences in the gamma band (31-45 Hz) for selected names in frontocentral electrodes from 300 ms. Again, the effect for faces did not survive cluster correction.

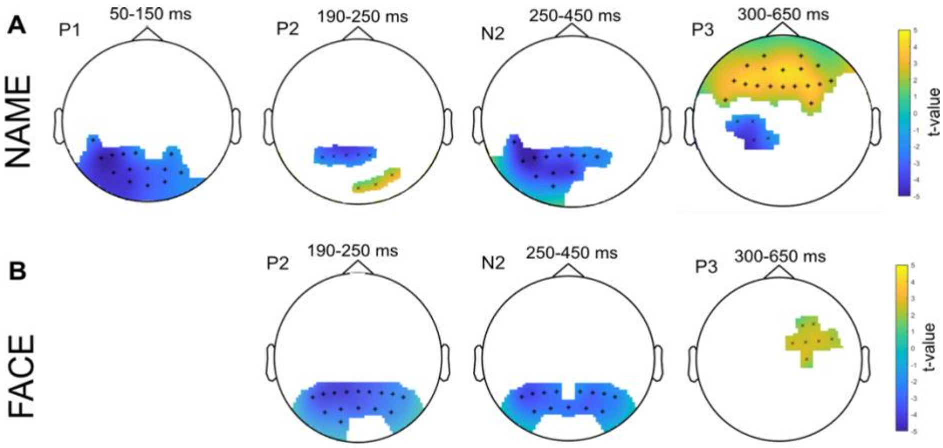


Figure 5.1. Event-related differences across conditions. Full table can be found in the Supplementary materials. Cueing effects for selected vs. ignored stimuli in attention. Yellow = cued > uncued, blue = uncued > cued. Asterisks (*) represent significant clusters for *p*<0.01 and exes (x) for *p*<0.05, after cluster-based permutation analyses performed for averaged frequency bins. Only significant electrodes are shown. (A) Names (B) Faces.

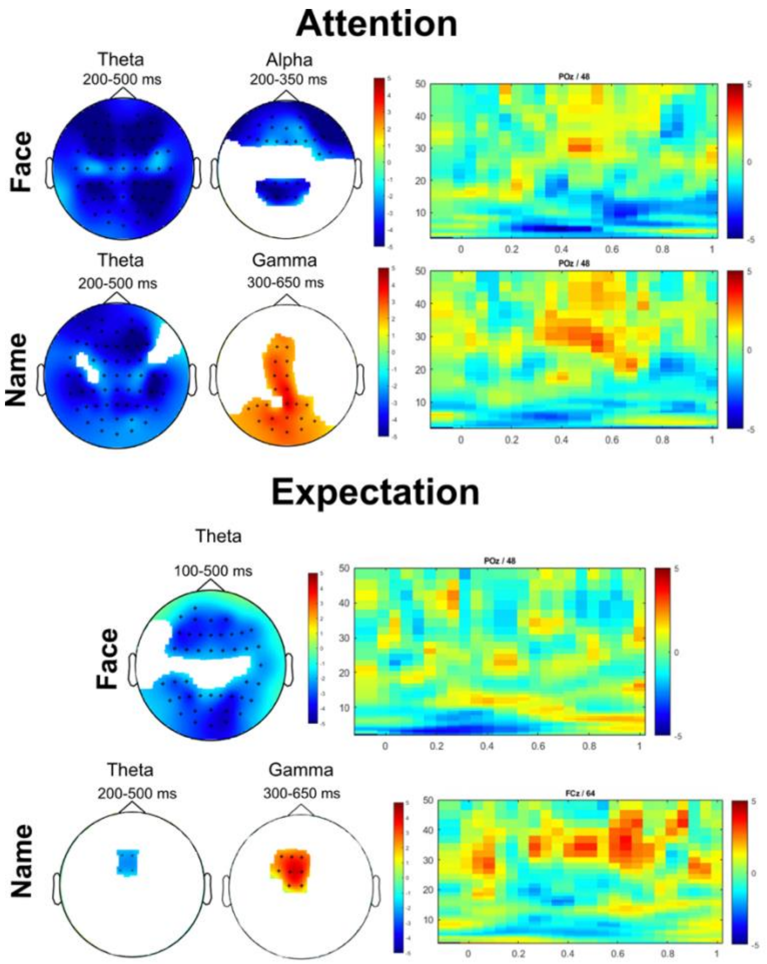


Figure 5.2. Oscillatory differences across conditions. Results show mass-univariate t-test comparisons between cued and uncued trials of attention and expectation, split by category (face or name). Topo-plots show the average of the significant window (indicated over the nose of each head). Time-frequency plots show t-values for all time-points, but only for one example channel. Red shows positive t-values (cued>uncued) while blue shows negative t-values (uncued>cued). Only significant electrodes are shown. Asterisks (*) represent significant clusters for $p<0.01$ and (x) for $p<0.05$, after cluster-based permutation analyses.

On the other hand, differences in expectation were more subtle. Uncued faces and names showed increased activity in the Theta band between 200 and 500 ms from target onset, the former in frontal, frontocentral and occipital electrodes (t(sum) = -666.43, $p<0.001$), and the latter in 4 fronto-central electrodes (t(sum) = -8.55, p = 0.047). Similar to attention, cued names induced larger gamma band activity from 300 ms to 600 ms in central electrodes (t(sum) = 48.57, $p<0.001$).

5.3.2.2. Probable stimuli reduce target decoding

To analyze prior based perception of targets, we classified faces and names in attention and expectation in both cued and uncued trials. All conditions were significant during at least one second after stimulus onset (Figure 5.3). The time resolved ANOVA with factors Block and Cueing showed several differences across decoding conditions. Main effects of block and cueing were significant for ~100 ms at the peak decoding accuracies. Although there were no significant interaction clusters, planned, within-block comparisons revealed that expected targets showed decreased decoding, in accordance with the dampening hypothesis. Note that this cannot be explained by

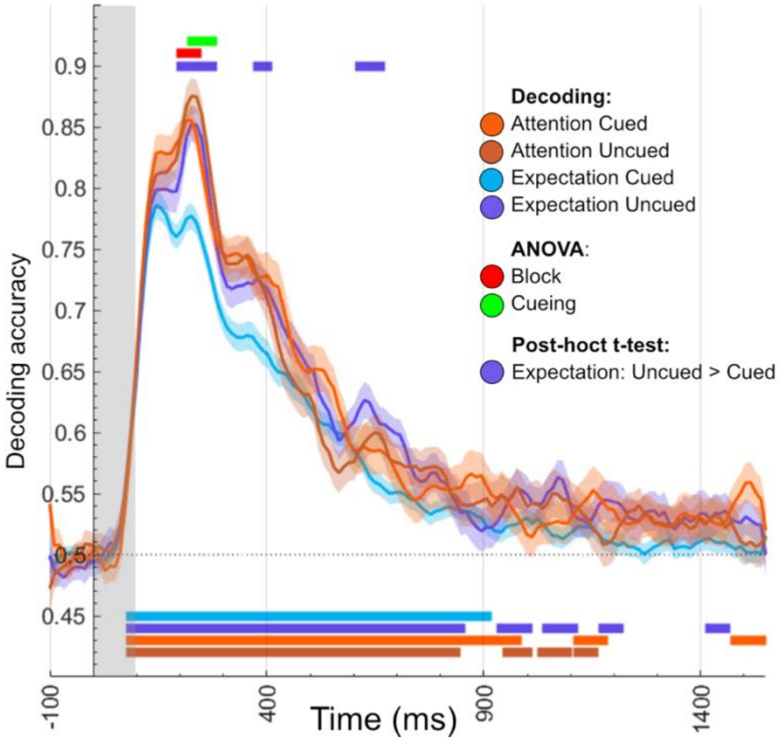


Figure 5.3. Time-resolved target decoding results. Faces vs. names target classification split by block and cueing conditions. Attention is represented in orange, and expectation in blue. Lighter tones indicate cued trials (light blue and orange) and darker tones indicate uncued (dark blue and brown). Horizontal colored lines below the chance line indicate statistical significance against chance within each block. Horizontal lines over the accuracy results indicate the results of the ANOVA (red = block, green = cueing). The dark blue horizontal in the top of the figure show the intervals where planned t-tests showed better decoding for uncued vs. cued trials of expectation. Gray shading indicates target presence onscreen.

different number of observations, since all conditions were subsampled to the less populated (trial-wise) condition. Attention showed no significant differences.

TGMs are depicted in Figure 5.4. We performed the decoding in each of the four conditions (block x cueing), plus an ANOVA with factors block and cueing, and planned *t*-tests in a similar fashion to the previous analysis (see Methods). This renders 9 different matrices shown in Figure 5.4 (4 conditions + 3 ANOVA factors + 2 planned *t*-tests). To help the reader, when we mention a specific matrix in text, we will specify line and column coordinates in the figure (e.g. ANOVA block effect is in line 3, column 1). Decoding analyses showed that the underlying patterns generalized across most of the analysis window in the four conditions (lines 1 and 2, columns 1 and 2). The main effect of block (3, 1) was present in a small cluster that matched the peak of classification decoding in the diagonal at 150 ms. The interaction effect between block and cueing (3, 3) did survive off the diagonal, again near peak latencies. Post-hoc *t*-tests explained this result by showing better decoding of cued targets in attention at 500 ms (1, 3), and a larger effect for uncued expectation targets

(2, 3), that were better decoded during most of the interval, inside and outside the diagonal.

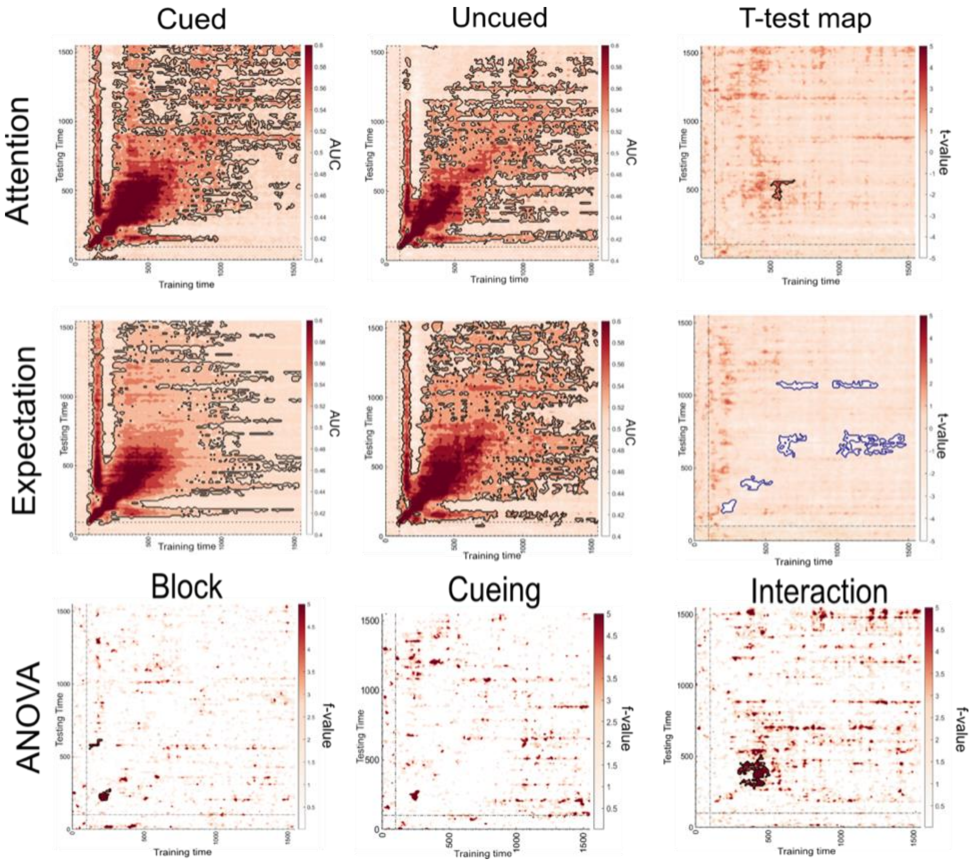


Figure 5.4. Temporal generalization results of target-locked analyses. Matrices show the result of training in every time-point (X axis) and testing in every other time-point (Y axis). Note that the diagonal of this plot matches the results in Figure 3. Dashed boxes in both axes show the window of the target onscreen. Significant positive clusters are outlined in black. The nine matrices are specified as line (1 to 3) and column (1 to 3), as in (row, column). Temporal generalization matrices for the four main conditions are: attention cued (1,1), attention uncued (1,2), expectation cued (2,1) and expectation uncued (2,2). The F values of the three ANOVA factors are in row 3: Block (3,1), Cueing (3,2) and Interaction (3,3). T-values of the planned t-test between cueing conditions of the same block are in column 3: attention (1,3) and expectation (2,3). Blue outline depicts significant negative clusters. Red indicates positive t-values (cued>uncued) while white shows negative t-values (uncued>cued).

5.3.2.3. Probable stimuli generalize from target to cue, relevant stimuli do not

To explore whether the reduced accuracies found in cued expectation trials reflected an attenuated and less stimulus specific form of processing, we performed a face vs. word classification using the complete window of analysis, locking trials to the cue onset (see Chapter 3 Methods for the cross-classification design). Since we were interested in how activity generalized from cue to target, we directly performed temporal generalization analyses (Figure 5.5). This resulted in four sections of generalization (only cue, only target, train cue and train target), described in the Methods and outlined in Figure 5.5A. As in the previous section, references to the results figure specify the TGMs with (line, column).

The only-cue section (see Figure 5.5A) was studied to evaluate whether results replicated the findings of Peñalver et al. (2023) even with the reduced power associated with splitting each condition in cued and uncued trials and subsampling. The four conditions (1:2, 1;2) showed evidence of anticipated decoding, although the results were reduced compared to previous reports. The two-way ANOVA did not reveal an effect of block (3,1) or an interaction (3,3) between block and cueing. A main effect of cueing was found in a small cluster at 500 ms (3,2), although since conditions were equated in the number of trials and there should not be any differences in cueing before target presentation, this result is likely spurious.

In the only-target section we sought to study the specificity of representations. This is, by training and testing in different trials split in two parts (remember the cross-classification design, see Methods), we expected to find more generalization if representations are less trial-specific (dampening) and worse accuracies if they are more trial-specific (sharpening). The results showed clear evidence of generalization in all four conditions, although there were differences across the different conditions. The ANOVA revealed a main effect of block (1,3), that was sustained during the first second after target onset, mostly outside the diagonal. This effect reflected an overall better generalization in expectation trials compared to attention. Cueing (3,2) and the interaction (3,3) effects were not significant. Crucially, planned comparisons showed several clusters where uncued attention trials generalized better than cued ones (1,3). On the other hand, generalization in expectation was similar between cued and uncued trials (2,3).

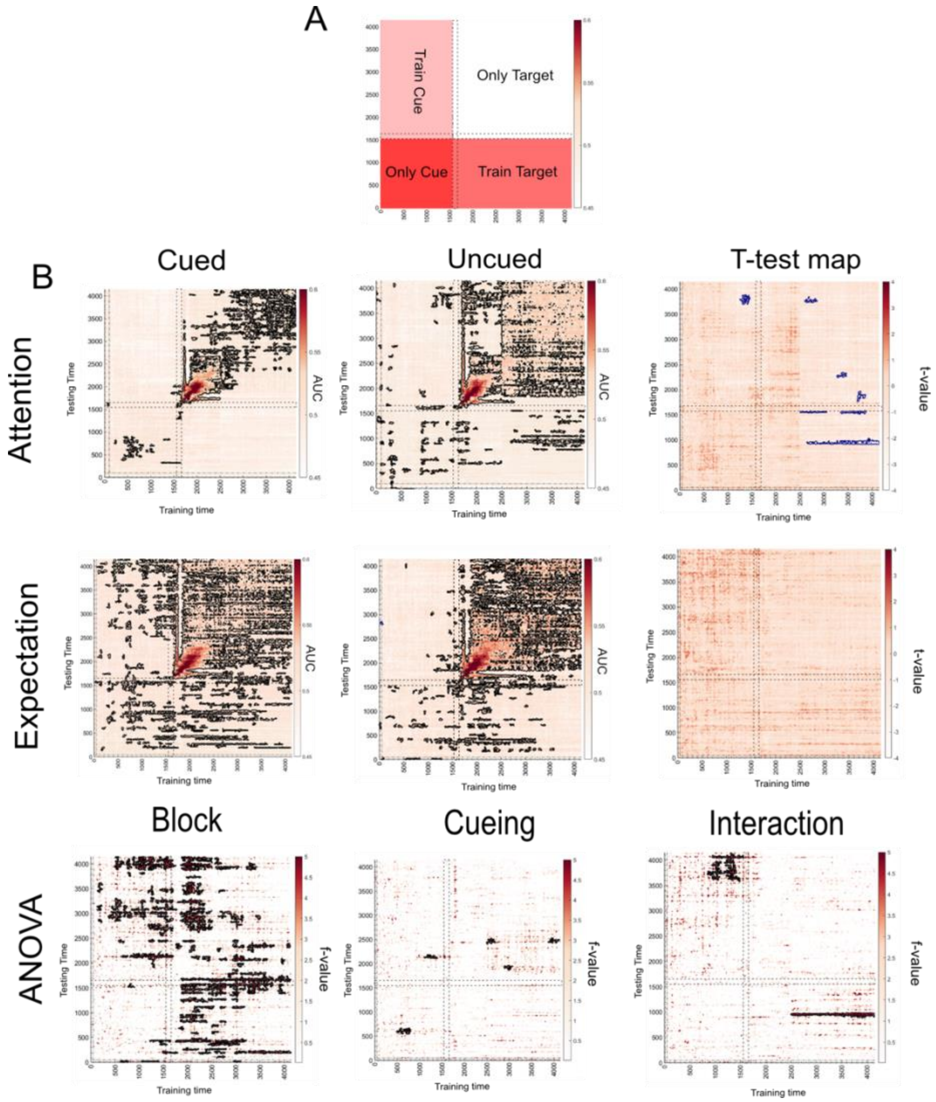


Figure 5.5. Temporal generalization results of target-locked analyses. A) Schematic description of the TGMs subdivisions. B) Matrices showing the result of training in every time-point (X axis) and testing in every other time-point (Y axis). Dashed boxes in both axes show the window of the cue (right by the axes, 50 ms windows) and target (from 1550 in both axes) onscreen. Significant clusters are outlined in black. In post-hoc t-test plots, black outlines mean significant cued>uncued, and blue outlines indicate significant uncued<cued. Figures are ordered in the same way as Figure 5.4.

The train cue and train target sections of the matrices were studied to examine the similarity between cue and target representations, as a correlate of the specificity of representations. We trained in different points of the cue interval (anticipate faces vs. names), and tested in different points of the target (see faces vs. names), and vice versa. If stimulus processing is less specific (dampening), it should be more similar to the cue and hence show larger cross-decoding accuracies. Conversely, if representations are sharpened, they should be more trial specific and hence show worse generalization between cue and target. Since we did not have separate specific hypotheses regarding the direction of the classification (train cue, or train target), and results were similar, we describe these two sections together. Uncued attention (1,2) trials showed significant decoding only between 1200-1500 ms after cue onset and 2000 ms after target onset, but cued trials (1,1) did not show any evidence of cue-target cross-decoding. On the other hand, cued expectation (2,1) trials showed large evidence of generalization between cue and target during the whole window. Finally, uncued expectation trials (2,2) only showed evidence of generalization target to cue, but scarce results in the opposite direction. The ANOVA indicated a large effect of block (3,1), with expectation showing larger evidence of generalization. Cueing (3,2) was only significant 1200 ms after cue onset, while the interaction (3,3) showed a cluster at the end of the target presentation window. Planned comparisons showed significantly lower cross-decoding for selected targets (1,3). In the expectation condition, despite the striking difference in significant over-chance decoding, differences across conditions did not survive cluster correction (2,3).

5.3.4. Discussion

In this first experiment, time-resolved analyses of prior based anticipation during visual perception in contexts of selective attention and perceptual expectations provided evidence in favor of the dampening hypothesis in expectation, alongside evidence that suggested sharpened representations for attention.

Univariate analyses returned a puzzling set of findings. The ERPs that tracked the effects of expectation suppression and attentional enhancement of perceptual processing, showed mixed results. Contrary to our hypotheses, we did not find increased amplitudes for perceptual ERPs for both cued attention and uncued expectation stimuli. Visual potentials (P1, P2, N2) showed larger amplitudes for uncued than cued attention targets, which goes in opposite direction than previous literature (e.g. He et al., 2004; Störmer & Alvarez, 2014). The reasons for this inversion are unclear, being perhaps related to the nature of the paradigm used. This will be discussed in detail in the general discussion. Modulation of the P3 did agree with previous results, as selected targets generated increased amplitudes in this potential, in line with previous results that link the P3 with selection of relevant targets (Potts, 2004).

Time-frequency decompositions results were closer to previous literature. Uncued attention trials consistently displayed higher theta power, which may be related to the change in response demands associated with this condition, and proposed as potential explanation for the early ERP results. Theta increases have been related to increased control demands (Cavanagh & Frank, 2014). On the other hand, alpha power was also larger for uncued trials, in line with research that proposes that alpha oscillatory activity reflects reduced attention in occipital and central positions (Bagherzadeh et al., 2020; Feng et al., 2017). Crucially, gamma was larger in occipital electrodes for cued trials, possibly reflecting the increased neural firing rate associated with increased activity for attended stimuli (Doesburg et al., 2008; Gruber et al., 1999).

To investigate potential effects of dampening or sharpening across conditions, we employed MVPA under the assumption that sharpening would be associated with increased fidelity (accuracy) of the neural codes, while dampening should result in lower decoding accuracy (de Lange et al., 2018; Kok, Jehee, et al., 2012). As expected, decoding of faces vs. words was significant in all block x cueing conditions. Although there were no differences between cued vs. uncued trials in attention, there was a clear effect in expectation, where unexpected trials were better decoded from 150 to 700 ms. Importantly, this result in expectation generalized outside the diagonal in several clusters between 500 and 1500 ms, suggestive of a "dampening" effect that was stable and lasted throughout most of the trial, indicating that the suppression of redundant neural activity is maintained from target perception.

Dampening accounts argue that the effect leads to reduced representations that are less specific (Press et al., 2020), which resonates with proposals of neural anticipations that understand preparation as a reduced form of perception (Battistoni et al., 2017; Koenig-Robert & Pearson, 2021). Hence, dampened representations during target processing should cross-generalize better to different stimuli of the same category, and to the cue interval. To test this, we studied the generalization between cue-locked and target-locked representations by performing a cross-decoding analysis, training and testing in different trials. Importantly, in this analysis we implemented a cross-classification that avoided perceptual confounds during the anticipation window, which trained and tested in different halves of the dataset. In cued attention trials we found little cross-classification compared with uncued trials, which showed larger clusters near the end of the cue interval. This result is compatible with sharpened target processing, since activity could be more specific to each trial and hence less similar to anticipation. Attentional selection has been shown to sharpen sensory representations repeatedly

(Goddard et al., 2022; Vaziri-Pashkam & Xu, 2017; Woolgar et al., 2015). Conversely, expected trials were the only ones to generalize to the cue window, in line with the dampening hypothesis (Blank & Davis, 2016; Kumar et al., 2017; Walsh & McGovern, 2018). However, expected and unexpected trials were not significantly different between them, so the conclusions of this analysis must be taken with caution.

Notably, during the "only target" window, there were significant off-diagonal clusters where unattended trials were better decoded than attended ones, as well as an important effect of block and (less pronounced) interaction. If attention sharpens neural coding, representations could become more stimulus specific and would show worse cross-decoding across different stimuli of the same category, supporting the cue-target cross-decoding findings. However, we did not have a priori hypotheses regarding this effect, so our interpretation of this interesting result is speculative, and needs further support.

Altogether, the novel use of EEG and temporal generalization to study these effects provides evidence that both sharpening of attended stimuli and dampening of unattended is temporally stable. However, since this technique is agnostic regarding the brain regions that can be involved in the effects, we studied the involvement of frontoparietal and visual regions using fMRI in Experiment 2.

5.4. Experiment 2

5.4.1. Methods

Participants, procedure and data acquisition

The dataset used in this section is the same as the one used in Chapter 4 of the present manuscript. Subjects, apparatus, stimuli, procedure, data acquisition and preprocessing steps are described at length in that section.

ROI definition

Regions of Interest (ROIs) were selected based on a priori hypotheses guided by previous studies of dampening and sharpening effects (Kok & de Lange, 2015; Richter et al., 2018). Specifically, we focused on visual and MD regions, bilaterally (8 in total, 4 of each). Visual regions were V1, VVC, FG, and LOC. MD regions included IPS, AI/FO, IFS and ACC/pre-SMA. All ROIs were obtained from the Human Connectome Project multimodal parcellation map (HCP-MMP1.0, Glasser et al., 2016), were extracted using Matlab code (https://github.com/davidwisniewski/fmri-extract-HCP-mask) and were registered back to each participant's native space using the inverse deformation fields obtained during segmentation.

Analyses

General Linear Model (GLM)

A GLM was performed to estimate activity changes across conditions, and to obtain the beta images to be used in subsequent multivariate analyses. We included cue and target regressors in the model, although here we only focused on target activity. Cue regressors were modeled using presentation time (500 ms) plus the subsequent jittered cue-target interval for each trial. They were divided by Block (attention and expectation), Category prediction (Faces and Names) and Cue shape (Shape 1 and Shape 2). Importantly, although there was a total of 8 different cue regressors, they were distributed across different runs. That is, in one run/block, cues could only be of a particular condition (e.g. Expectation), and the cues predicting a certain category had the same shape during the run (e.g. Face-shape 1 and Name-shape 2). Hence, each run included only 2 cue regressors. Target information was modeled using the presentation time of the stimulus on screen (500 ms), and consisted on regressors for the conditions of Block type (attention and expectation), Category (faces and names) and Cueing (cued vs. uncued). Again, although there were 8 different target regressors, they belonged to particular runs, which resulted in runs having 4 target regressors each. Crucially, conditions are by design unbalanced in number of trials, since cued trials in expectation were three times more common than uncued trials. This means that betas of this condition are likely to have a higher signal-to-noise ratio, which in turn would likely bias decoding results. To avoid this issue, before implementing the GLM, all conditions (attention and expectation, Cued and Uncued) were subsampled to the number of trials of the less populated condition (uncued expectation). Specifically, for each run, we identified the smaller number of trials of all conditions, and then randomly chose the same number of trials in the remaining ones. Despite the loss of power, this approach made possible to directly compare decoding results across conditions. Finally, the model also included as regressors movement information of each participant, obtained during realignment. All regressors were convolved using the canonical hemodynamic response function (HRF).

Univariate

Univariate comparisons in individual ROIs were performed on the beta images of the GLM. There were two main contrasts of interest: attention Cued vs. Attention Uncued, and expectation Cued vs. expectation Uncued. For ROI analyses we obtained the beta values of each voxel and condition and then averaged across all voxels, leaving one mean beta value per ROI, condition and participant. These were entered in a two-way repeated measures ANOVA with factors Block (attention vs. expectation) and Cueing (Cued vs. Uncued). Moreover, specific a priori hypotheses regarding effects of expectations and selection led to planned comparisons between cued and uncued trials separately for attention and expectation, using a repeated measures *t*-test. We used FDR to correct for multiple comparisons across ROIs.

For whole brain univariate analyses the whole brain image of each participant and contrast was entered in a second-level analysis, where they were compared using a *t*-test. Significance was stablished by first selecting voxels that passed a threshold of $p<0.001$, and a subsequent cluster size corresponding with $p<0.05$, FWE-corrected. This method has been shown to appropriately correct for multiple comparisons (Eklund et al., 2016).

Multivariate

Decoding was based on a two-class classification approach, using beta images from the GLM with The Decoding Toolbox (v 3.999F). We trained a classifier (Support Vector Machine, SVM) with a leave-one-run-out cross-validation, training in all but one run, and then testing on the remaining one. This was repeated across all runs, and the results averaged. To ensure an unbiased classification we report balanced accuracies. With this approach, we performed four decoding analyses in which we classified faces vs. names, for the conditions: attention cued, attention uncued, expectation cued, and expectation uncued. This returned one accuracy value per subject, ROI and condition. Next, to investigate whether decoding accuracies were higher in either condition we entered the four condition results into a two-way repeated measures ANOVA with factors Block and Cueing. We used FDR to correct for multiple comparisons across ROIs.

We performed an exploratory whole-brain analysis to reveal additional regions that might have reflected dampening or sharpening effects. We used a searchlight procedure to assess classification values between the two classes (spheres of 4 radii, 251 voxels). Afterwards, the resulting images were registered to the MNI space and smoothed. Group statistics applied a one-sample *t*-test against zero. We used the same method described in univariate analyses to correct for multiple comparisons. To investigate whether decoding accuracies were higher in either condition of the Cued-Uncued pairs, separately for each block, we calculated the difference between decoding results for each voxel. With the resulting images of all participants, we performed one-sample *t*-tests, and established significance with the same procedure described above.

5.4.2. Results

5.4.2.1. Univariate cueing differences

ROI

To reveal univariate differences in activation associated with cueing in the MD, we analyzed four regions that have been shown to exhibit differences in previous literature. All frontoparietal regions showed significant activations compared to the baseline (all $p<0.001$), as evidence by a one sample t-test (see Supplementary Table 5.3). However, two-way ANOVAs in each ROI showed no overall differences in any regions for any factor (all $p>0.05$, see Figure 5.6, Table 5.1 and Supplementary Table 5.5).

All visual regions showed significant activations compared to the baseline (see Supplementary Table 5.4). The ANOVA revealed that there was no main difference between attention and expectation (see Figure 5.7 and Table 5.2). There was, however, a main effect of cueing in VVC, FG and LOC. Post-hoc comparisons revealed that in these ROIs cued trials showed higher beta values in both attention and expectation. Although the interaction was not significant (all $p>0.09$), we performed a-priori planned comparisons to explore the specific effects in each block (Supplementary Table 5.6). We found significant differences only in attention, where cued trials showed significantly higher beta values in the VVC, while the others did not survive FDR correction.

Table 5.1. Univariate ROI ANOVA, MD regions.

ROI	Factor	F	p	ηp2
ACC/pre-SMA	Block	0.102	0.751	0.002
	Cueing	0.136	0.714	0.003
	Interaction	2.031	0.161	0.045
AI/FO	Block	0.059	0.809	0.001
	Cueing	2.142	0.151	0.047
	Interaction	1.141	0.291	0.026
IFS	Block	0.135	0.715	0.003
	Cueing	0.121	0.729	0.003
	Interaction	0.011	0.916	0.001
IPC	Block	0.969	0.331	0.022
	Cueing	0.894	0.350	0.020
	Interaction	2.910	0.095	0.063

Note: The term interaction refers to the interaction between block and cueing.

Table 5.2. Univariate ROI ANOVA, visual regions.

ROI	Factor	F	p	ηp2
V1	Block	2.232	0.143	0.049
	Cueing	4.383	0.435	0.014
	Interaction	1.299	0.924	0.001
FG	Block	1.061	0.309	0.024
	Cueing	9.367	0.004*	0.179
	Interaction	1.030	0.316	0.023
VVC	Block	1.739	0.194	0.039
	Cueing	10.616	0.002*	0.198
	Interaction	2.886	0.097	0.063
LOC	Block	1.508	0.226	0.034
	Cueing	6.893	0.012*	0.138
	Interaction	0.102	0.750	0.002

Note: The term interaction refers to the interaction between block and cueing.

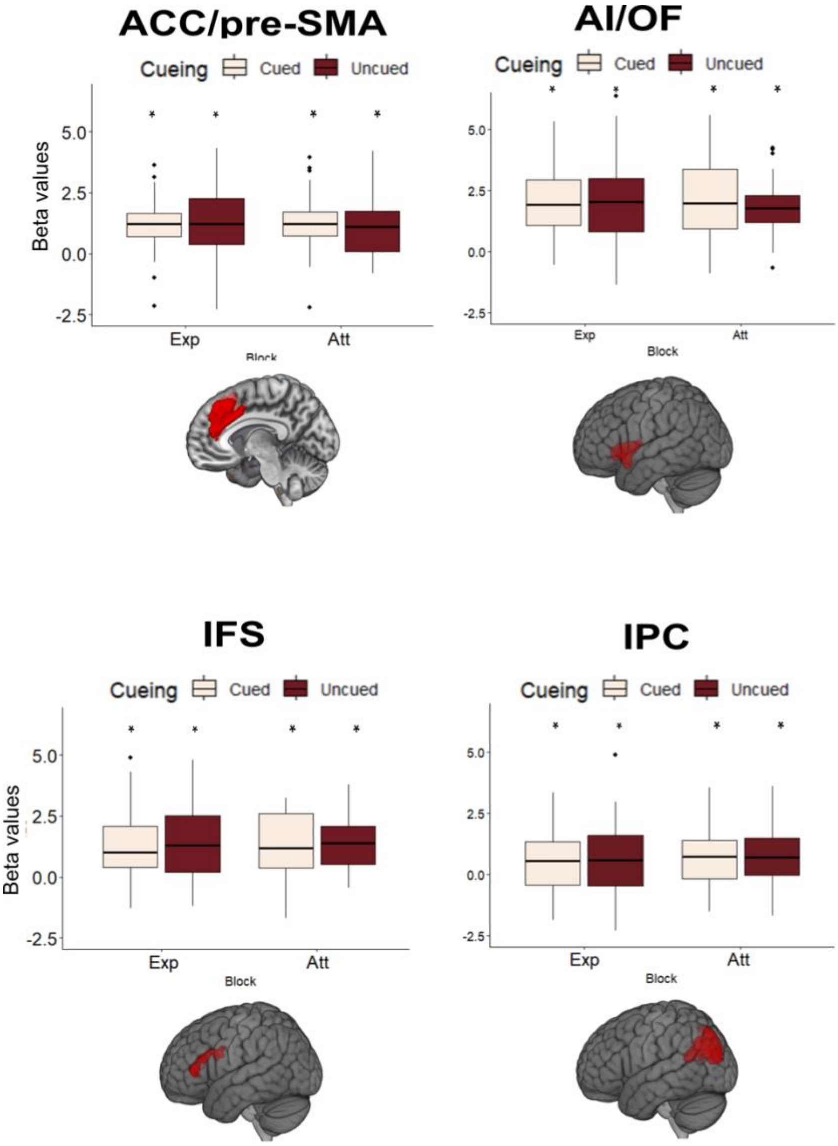


Figure 5.6. Univariate analyses results for MD ROIs. Boxplots show the differences in beta values between the different conditions. The horizontal black line inside boxes represents the median, while the limits of the box indicate the first and third quartile. Whiskers indicate the 1.5 inter quartile range for the upper and lower quartiles. ROIs are highlighted below each boxplot.

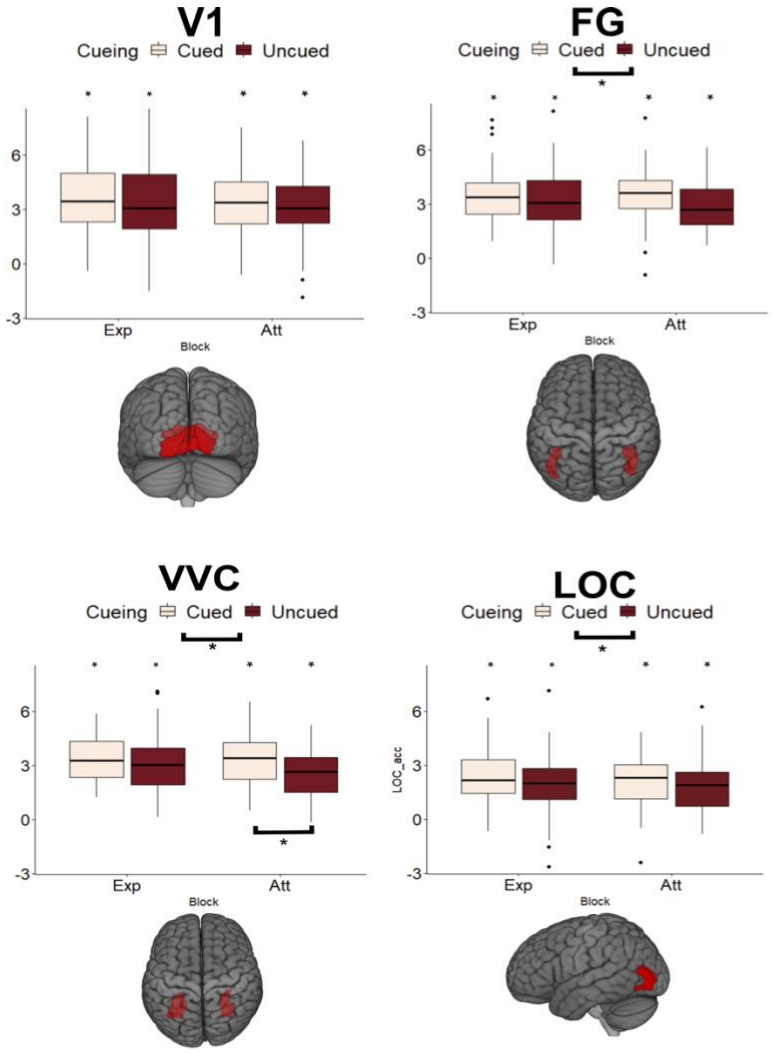


Figure 5.7. Univariate analyses results for visual ROIs. Boxplots show the differences in beta values between the different conditions. The horizontal black line inside boxes represents the median, while the limits of the box indicate the first and third quartile. Whiskers indicate the 1.5 inter quartile range for the upper and lower quartiles. Asterisks above two boxes depict significant differences between cueing. Black bars below the cued-uncued legend indicate a main effect of cueing. ROIs are highlighted below each boxplot.

Whole-brain

The whole-brain univariate comparisons across cueing instances of attention and expectation were performed to evaluate expectation suppression and attentional enhancement in regions outside the ones chosen a priori (Figure 5.8, Table 5.3). Cued attention trials led to increased activity in the Fusiform Gyrus (FG), the ACC and part of the Hippocampus, whereas uncued trials led to increases in regions often associated to the frontoparietal control network (FPN), such as the middle frontal gyrus and the inferior parietal sulcus (IPS). On the other hand, cued trials in expectation led to increases in perceptual regions such as the inferior temporal gyrus (ITG). Surprisingly, the reverse contrast did not show significant results in any region.

Table 5.3. Whole-brain Univariate Analyses Results.

Contrast	Region	Coordinates	Cluster size (k)	Z	p	T(peak)
Att. Cued > Uncued	FG Left	-40, -44, -14	791	5.39	<0.001	5.39
	HC Left	-30, -8, -32	711	4.97	<0.001	711
	ACC Right	2, 40, -18	456	4.37	0.002	456
	ACC Right	6, 10, 28	286	4.21	0.018	286
Att. Uncued > Cued	PMd	-30, 0, 58	1080	4.92	<0.001	5.76
	FP	-42, 50, -4	391	4.15	0.005	4.63
	IPS	-42, -56, 38	586	4.04	<0.001	4.48
	MFG	40, 28, 38	243	3.8	0.034	4.17
Exp. Cued > Uncued	ITG Right	46, -58, -10	238	4.31	0.032	4.85
	STG Right	60, -12, 4	222	3.88	0.40	4.27

FG = Fusiform gyrus, HC = Hippocampus, ACC = Anterior Cingulate Cortex, PMd = dorsal Premotor Cortex, FP = Frontal Pole, IPS = Intraparietal Sulcus, MFG = Middle Frontal Gyrus, ITG = Inferior Temporal Gyrus, STG = Superior Temporal Gyrus

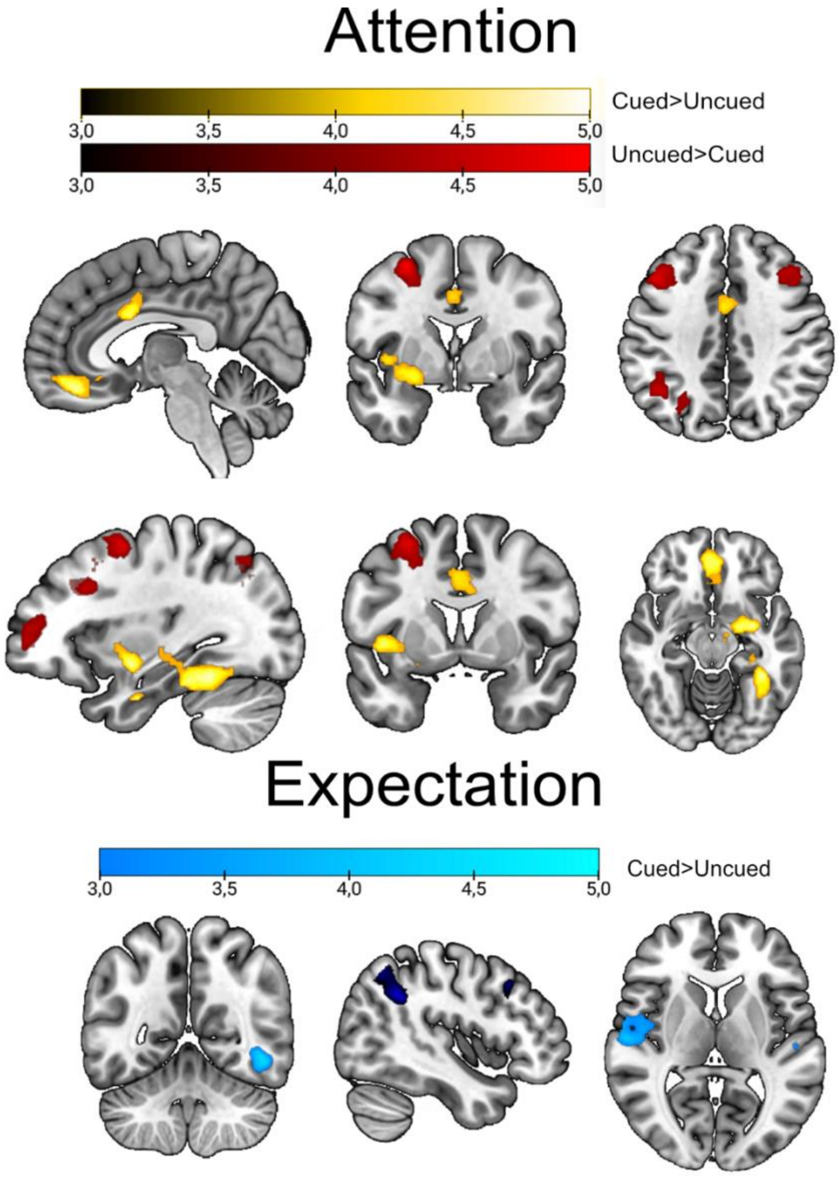


Figure 5.8. Mass univariate results. Contrast GLM results for cued vs. uncued trials in attention and expectation. Scales reflect t-values. Top shows results for attention. Golden areas highlight significant clusters for cued trials, while red indicates uncued trials. Bottom shows results for expectation. Only clusters of larger activity for cued trials are shown, since there were no significant uncued clusters.

5.4.2.2. Multivariate analyses

ROI

To study the extent to which relevant information could be decoded, and hence whether our results favor a model of dampening or sharpening in attention and expectation, we classified faces and names during target processing in MD and visual regions.

In MD ROIs, a one-sample *t*-test against zero of the decoding result minus chance (see Supplementary Table 5.7) for *t*-values and mean decoding accuracies) revealed that decoding was significant in IFS, in all conditions except uncued attention trials, and IPC, for uncued expectation and cued attention trials. All other regions and conditions did not show evidence of significant classification after FDR correction (p>0.012). We compared the decoding of the four conditions using a two-way repeated measures ANOVA did not reveal any significant differences in any of the ROIs and factors (Figure 5.9 and Table 5.4). Noteworthy, all regions exhibited the same trend of better decoding for cued than uncued attention trials, and worse decoding for cued than uncued expectation targets (see Supplementary Table 5.9).

Table 5.4. Decoding ROI ANOVA, MD ROIs.

ROI	Factor	F	p	$\eta p2$
ACC/pre-SMA	Block	1.151	0.289	0.026
	Cueing	0.339	0.563	0.008
	Interaction	1.523	0.224	0.034
AI/FO	Block	3.639	0.063	0.078
	Cueing	0.410	0.525	0.009
	Interaction	0.896	0.349	0.020
IFS	Block	0.969	0.331	0.022
	Cueing	0.894	0.350	0.020
	Interaction	2.910	0.095	0.063
IPC	Block	3.371	0.073	0.073
	Cueing	0.022	0.884	0.001
	Interaction	4.031	0.051	0.086

Note: *p* FDR corrected for multiple comparisons = 0.025.

Table 5.5. Decoding ROI ANOVA, visual ROIs.

ROI	Factor	F	p	$\eta p2$
V1	Block	0.376	0.543	0.009
	Cueing	7.170	0.010*	0.143
	Interaction	16.319	< .001*	0.275
FG	Block	2.579	0.116	0.057
	Cueing	14.962	< .001*	0.258
	Interaction	26.282	< .001*	0.379
VVC	Block	7.668	0.008*	0.151
	Cueing	10.542	0.002*	0.197
	Interaction	14.237	< .001*	0.249
LOC	Block	0.582	0.450	0.013
	Cueing	5.365	0.025*	0.111
	Interaction	10.779	0.002*	0.200

Note: *p* FDR corrected for multiple comparisons = 0.025

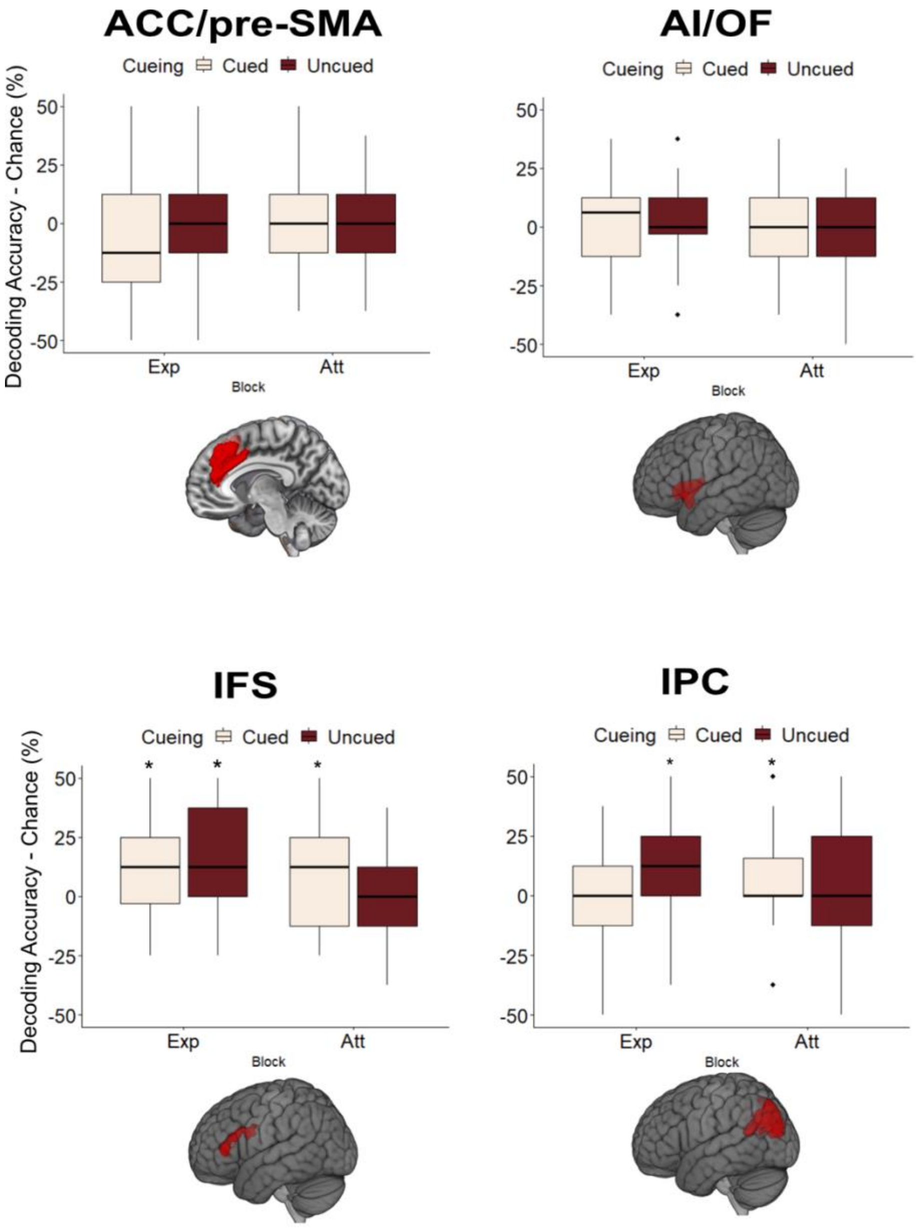


Figure 5.9. Decoding accuracies for each MD ROI and condition. The horizontal black line inside boxes represents the median, while the limits of the box indicate the first and third quartile. Whiskers indicate the 1.5 inter quartile range for the upper and lower quartiles. Asterisk above boxes indicate significant decoding of faces and names in that ROI and condition.

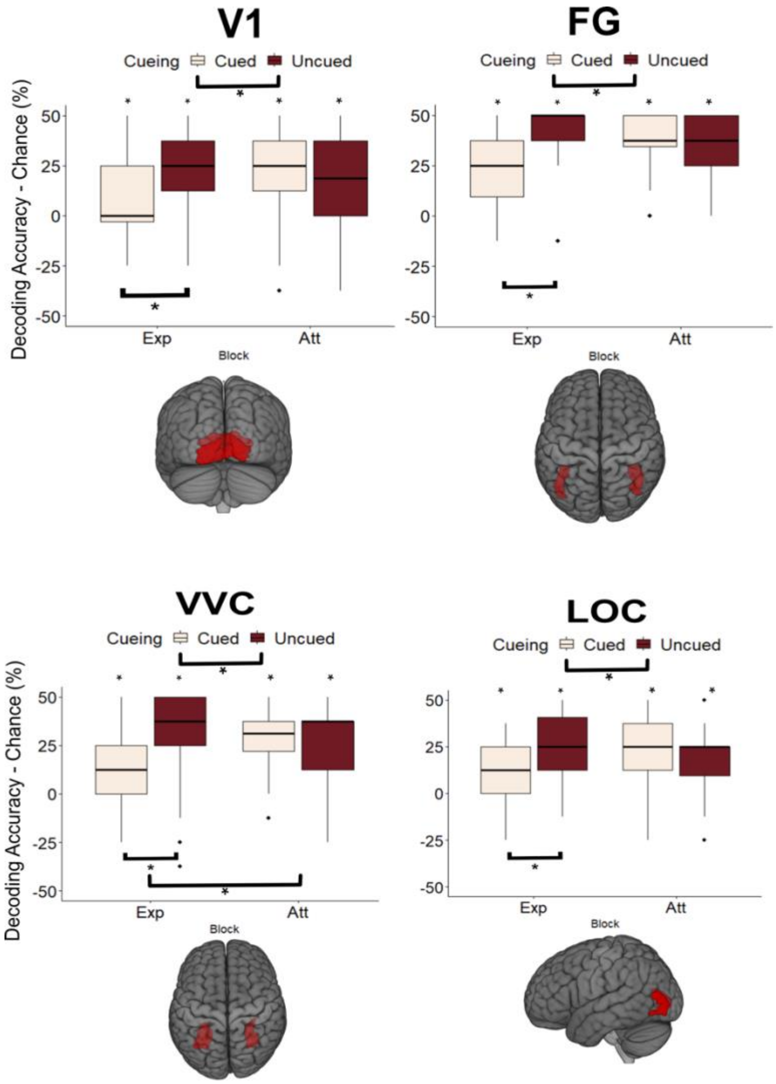


Figure 5.10. Decoding accuracies for each visual ROI and condition. The horizontal black line inside boxes represents the median, while the limits of the box indicate the first and third quartile. Whiskers indicate the 1.5 inter quartile range for the upper and lower quartiles. Asterisks above the cueing legend indicate a main effect of Cueing, asterisks over two pairs of boxes show a significant main effect of Block and asterisks over two boxes specify significance between two conditions.

We repeated the same procedure for visual ROIs. First, as expected, we found that all regions and conditions showed significant decoding of the target categories (Supplementary Table 5.8). When we compared the decoding results of the four conditions, the ANOVA (Table 5.5 and Figure 5.10) revealed a main effect of block on the VVC, as attention trials were overall better decoded than expectation ones. The main effect of Cueing was significant in all regions (all $p<0.025$), and in all of them it significantly favored uncued over cued trials. Crucially, the interaction was also significant in all regions (all $p<0.002$), with a stable pattern of better decoding for uncued vs. cued trials in expectation, and the opposite trend (albeit non-significant) in attention (Supplementary Table 5.10).

Whole brain

To further explore effects in regions other than the selected ROIs, we performed a target-locked classification of faces vs. names at the whole-brain level (Figure 5.11). Within-condition decoding revealed classification of the perceived target in all conditions, peaking at the FG in all conditions except uncued attention, which peaked at V1 (see Table 5.6). Besides, cued expectation trials had another cluster in the IFG, and uncued expectation trials showed a significant cluster in the retro-splenial cortex (RSC).

Crucially, uncued expectation trials were better decoded than expected ones in the same FG clusters and broadly the VVC (Table 5.7). There was also a cluster in the intraparietal sulcus. In contrast, there were no significant differences between cued and uncued trials in attention in either condition.

Table 5.6. Target-locked decoding results for cued and uncued trials in attention and expectation.

Contrast	Region	Coordinates	k	Z	p	Acc(peak)
Att. Cued	FG	38, -76, -8	26362	>8	<0.001	75,18%
Att. Uncued	V1	22, -96, 6	15397	>8	<0.001	75,24%
Exp. Cued	FG Right	40, -52, -14	3893	7.09	<0.001	71.27%
	IFG Left	-50, 10, 12	4002	6.08	<0.001	64,94%
Exp.Uncued	FG Left	-40, -56, -14	26618	>8	<0.001	83,77%
	IFG Left	-50, 2, 22	1246	4.56	0.005	60,86%
	RSC	0, -54, 24	853	4.41	0.019	60,46%

Note: *p* FWE corrected for multiple comparisons. k = cluster size. FG = Fusiform Gyrus, V1 = Primary visual cortex, IFG = Inferior frontal gyrus, RSC = Retro-splenial cortex.

Table 5.7. Difference between Cued and Uncued conditions. Only significant clusters are shown.

Contrast	Region	Coordinates	k	Z	p	Acc(peak)
Exp. Uncued-Cued	FG Left	-42, -62, -10	14215	6.65	<0.001	18,92%
	S1 Right	10, -44, 62	2105	4.92	<0.001	14,07%

S1 = Primary Somatosensory Cortex. Acc(peak) column shows the peak difference in accuracy percentage between uncued and cued trials.

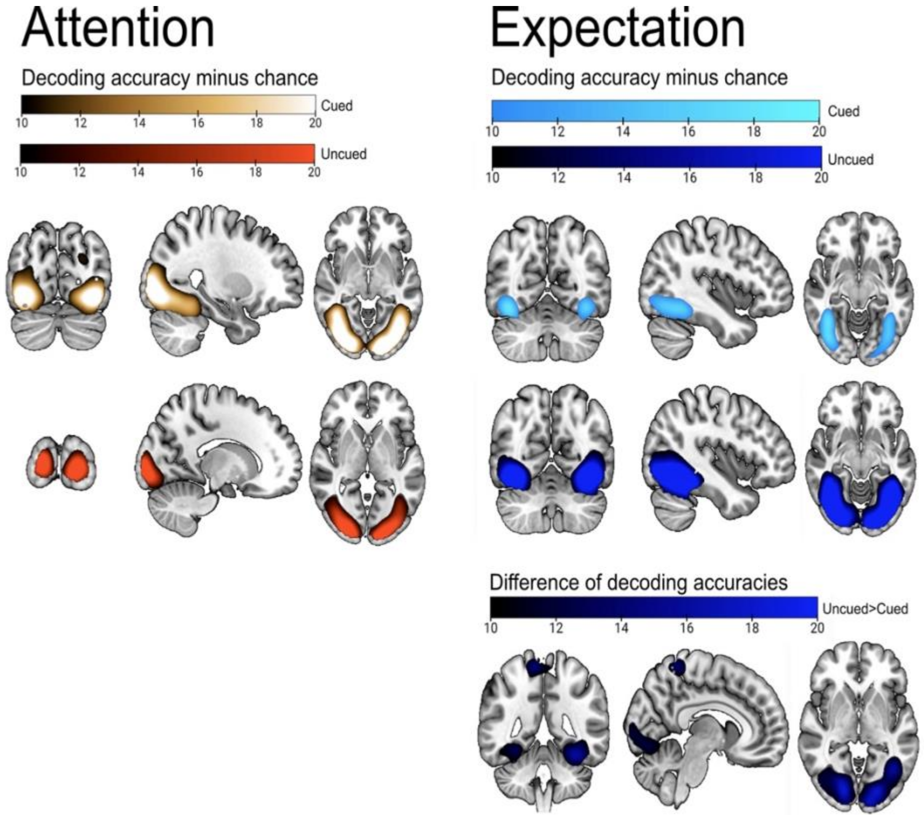


Figure 5.11. Target-locked decoding results. Top four rows show regions with significant over-chance decoding for Faces vs. Names targets in the four conditions.

5.4.3. Discussion

In this second part of the study, we investigated how attention and expectation cueing affect processing on frontoparietal and visual regions, and whether results suggest sharpening or dampening effects, with the same rationale followed for the EEG data. Employing the same paradigm, adapted to fMRI, the main findings indicate that overall cued expectation trials were decoded worse, supporting the dampening account, although here we did not find evidence for sharpened representations in cued attention targets.

To examine potential effects of attention enhancement (Kastner et al., 1999; Serences et al., 2004) and expectation suppression (e.g. Kok et al., 2012; Walsh & McGovern, 2018; Feuerriegel et al., 2021) we performed contrasts of univariate changes associated to each condition. We studied the effects in frontoparietal regions, such as those of the MD, since they have been shown to be involved in stimulus representation (Woolgar et al., 2015) and are involved in both attention and expectation contexts (Eger et al., 2007; Esterman & Yantis, 2010). Although all ROIs showed significant activations vs. the baseline during target processing, the ANOVA did not reveal any univariate differences between the four conditions. Other studies attempting to disentangle dampening and sharpening effects have reported similar results (González-García & He, 2021), while others that have attempted to decode information from these ROIs only report univariate differences in visual regions (Woolgar et al., 2015). Our results, however, imply that these regions are engaged during target processing in both attention and expectation.

In contrast, visual ROIs showed both significant activations vs. baseline during target processing as well as a cueing effect in all ROIs but V1. Overall, cued targets elicited increased activations, regardless of the relevance/probability manipulation. Although this is a frequent finding in the context of selective attention (e.g. Goddard et al., 2022; Grill-Spector et al., 2001; Kastner et al., 1998; Woolgar et al., 2015), we expected to find opposite results for expectation (Feuerriegel, Vogels, et al., 2021; Richter et al., 2018). Importantly, post-hoc tests were only signficant for attention enhancement, particularly in the VVC, which highlights the object-selective nature of the neural activations linked to selective attention in our task.

Univariate exploratory whole-brain searchlight analyses between cued and uncued trials, separately for each block, revealed several regions that were differentially affected by cueing in attention and expectation. In attention, cued trials showed increased BOLD in the hippocampus, the cingulate cortex, pre-SMA and VVC the ventral cortex (in line with the ROI results). The fact that the cingulate cortex appeared here but was not significant in ROI analyses could mean that the area we chose was too large and included voxels not particularly involved in stimulus representation. Notably, the hippocampus has been shown to be more active during encoding and retrieval of attended stimuli (Uncapher & Rugg, 2009), although by itself it does not appear to modulate selective attention (Rungratsameetaweemana et al., 2019). Uncued attention trials led to increased activity in IPS and some midle regions of the frontal cortex, mapping to the frontoparietal control network, which has been related to cognitive control and might reflect the change in response mappings associated with uncued attention trials. On the other hand, in expectation, in line with ROI results, cued trials showed higher activation levels in the inferior and superior temporal gyri, linked to processing of complex stimuli (Kumar et al., 2017; Mechelli et al., 2000).

To compare how cueing affected the fidelity of neural representations and compare between dampening and sharpening in contexts of relevance and probability, we used multivariate decoding analyses (e.g. Goddard et al., 2022; Kok et al., 2012; Walsh & McGovern, 2018; Woolgar et al., 2015). We did this by training a classifier to decode faces and names, separately for each cueing and block condition. ROI analyses showed that the target's category could be decoded in specific conditions in two frontoparietal regions, IFS and IPC. This was only the case for cued attention trials, a result that is is similar to the one obtained by Woolgar et al. (2015), with significant decoding for attended trials, but not unnatended. However, we did not replicate it for ACC/pre-SMA and AI/FO cortices, which, as in their experiment, might need higher stimulus complexity to become directly involved. Moreover, both predictive coding models and attention literature that manipulates probability have proposed FPN to be involved in contexts of expectation, which would be supported by significant decoding in both cued and uncued trials of expectation in the IFS. However, neither the ANOVA or the post-hoc tests revealed significant differences due to cueing in any of the conditions or between blocks, although there was a consistent trend in all frontoparietal ROIs for better decoding in cued vs. uncued attention trials, and the opposite effect in expectation.

However, decoding perceptual categories in these ROIs gave small accuracies compared to visual ROIs, which might make it difficult to find signficant results.

Visual ROIs showed very clear differences in expectation, but not in attention. This result was further confirmed by the searchlight analysis. Probable trials showed decreased classification accuracy in all visual regions, which gives support for the dampening hypothesis, replicating similar findings (Blank & Davis, 2016; Kumar et al., 2017; Walsh & McGovern, 2018). Although this interaction was significant in all ROIs, cued attention trials were not significantly better decoded than uncued ones.

5.5. General Discussion

Overall, the results employing two different neuroimaging techniques consistently show decreased decoding accuracies for cued expectation trials, providing strong support for the dampening hypothesis in expectation, and a lack of generalization between cued attention targets and preparatory cues, which provide tentative evidence of sharpened representations for selected stimuli.

One main solid finding across EEG and fMRI results is the lack of expectation suppression effects, as univariate activations for cued trials were not reduced. Although the study of dampening vs. sharpening originated to account for such univariate effect within the predictive coding framework (Friston, 2005; Kok, Jehee, et al., 2012), the effects have also be studied in the absence of univariate changes (González-García & He, 2021). It has been suggested that not all paradigms involving probability manipulations are equally able to capture expectation suppression, and this effect also depends on imaging methodologies (Feuerriegel, Vogels, et al., 2021). For instance, some the paradigms that report larger voltages for attended stimuli, manipulate attention by cueing the probability of incoming stimuli (e.g. He et al.), which might create confounds between attended and probable stimuli. In addition, our targets were not surrounded by other stimuli, in contrast with previous reports (e.g. Störmer & Alvarez, 2014). Specifically, cue-target paradigms in human electrophysiology frequently report an absence of this effect (Feuerriegel et al., 2018; Hall et al., 2018; Summerfield et al., 2011). Moreover, conditions with less trials tend to be associated with artificially inflated amplitudes (Thomas et al., 2004), which could be the case for the unexpected condition in some EEG studies reporting increased amplitudes in the unexpected condition (e.g. Tang et al., 2018). Importantly also, experiments that report effects of expectation suppression usually rely on a small set of images (e.g. Richter et al., 2018; Walsh & McGovern, 2018), which allows for stronger cue-target pairings and thus stronger learning of predictability. The current experiment used different images on each trial and thus induced expectations at a more abstract categorical level, which is different from an item-level prediction. This prevents precise stimulus expectation effects, since expected trials are linked to categories rather than specific, repetitive stimuli.

Another puzzling set of results is the difference between univariate EEG and fMRI results. While EEG data showed overall larger amplitudes for the uncued conditions, the effects were reversed in fMRI. Although some instances of previous research have successfully linked ERPs and fMRI activations (Corrigan et al., 2009; Natale et al., 2006), our data did not return converging results. However, since we used a similar paradigm in two methodologies with big differences in temporal precision, results might have been affected. While EEG trials lasted barely 4 seconds each and we could obtain measures during stimulus perception, easily separated from the preparation window and response times, the BOLD signal is slow, which was accounted for by using trials of up to 20 seconds, which might have changed stimulus processing to one experiment to the other. On the other hand, oscillatory effects and fMRI activations were potentially a more consistent match, specifically for attention. One set of mechanisms could explain this compatibility. The specific perceptual characteristics linked to gender may have been ignored in uncued trials, reflected by Alpha, and the switch to a different response set, could be indicated by increased Theta power. This effects on slower oscillations could be overlooked by fMRI contrasts, which are less sensitive to slow oscillations (Scheeringa et al., 2011). Conversely, cued trials would receive increased perceptual processing, marked by Gamma power. This agrees with the increased activations found for cued attention trials in perceptual regions, since the BOLD signal is particularly sensitive to gamma power (Magri et al., 2012).

In expectation we found converging evidence of dampened representations for probable stimuli, similarly to previous reports in fMRI (Blank & Davis, 2016), and monkey multi-unit recordings (Kumar et al., 2017). Moreover, fMRI results showed this effect for visual regions and, accordingly, EEG showed early temporal latencies. Predicted information, anticipated by cues, could be partially represented in perceptual regions once the target arrives (Peñalver et al., 2023), so redundant data could be filtered out to speed sensory processing, thus increasing behavioral accuracy. The potential alternative explanation of increased attention to unexpected trials (Alink & Blank, 2021), which could imply sharpened coding for uncued expectation trials is not supported by our data. First, this explanation would predict that probable stimuli should be decoded similarly than unattended ones. However, out of the four conditions, expected trials were specifically poorly decoded compared to all other conditions in the two experiments, suggesting that dampened representations are an effect which is specific for probability contexts. Second, cued trials in both attention and expectation showed increased univariate activations in visual regions, as well as increased gamma power. Hence, while increased activations for cued in expectation trials were associated with poorer classification accuracy in expectation, the same activations in attention were not linked to different decoding accuracies. Moreover, cue-target temporal generalization in EEG, and a non-significant trend in fMRI suggested the opposite effect. In the future, more powered studies can find stronger support for this diverging result.

In attention, we expected to find evidence of sharpened representations for selected stimuli, both in visual and frontoparietal regions (Woolgar et al., 2015). EEG data showed direct evidence for this effect in a small cluster during the temporal generalization analysis, while with fMRI comparisons across different ROIs or a whole-brain contrast did not return differences that passed significance levels. Moreover, EEG revealed that cued attention trials generalized worse to the preparatory period, in line with the assumptions of sharpening models (Desimone & Duncan, 1995). This implies that attended trials induce more detailed and specific representations, which might be associated to more appropriate and efficient responses. This effect could also be behind accounts that propose that selective attention increases the processing weights of attended stimuli, inducing more detailed representations (Feldman & Friston, 2010; Summerfield & Egner, 2014).

The current study is not without limitations. Clearly discriminating whether responses were dampened for expected stimuli, or sharpened for all other conditions is not possible without including a neutral control condition (Feuerriegel, Vogels, et al., 2021). Moreover, here we did not make a direct connection between the two experiments, but rather perform parallel analyses. Performing joint analysis could provide better insight into the spatiotemporal profile of dampening and sharpening effects. For instance, our design could be well suited to perform multimodal fusion (Cichy & Oliva, 2020), since the two experiments share the same paradigm adapted to exploit the advantages of each methodology.

Althogether, across two different imaging methodologies we provided evidence that probable stimuli undergo a “dampening” effect, while selected targets tend to elicit a sharpened, trial-specific representation of visual stimuli. This result was evident in visual regions, and extended over time. Moreover, we found that the relationship between univariate and multivariate activity is different in contexts of attention and probability. While attended stimuli elicit larger univariate activations, better decoding and worse cross-classification, expected stimuli showed similar larger univariate changes that were conversely related to poorer representations and increased across-trial and preparatory similarity. Overall, our results add to previous literature by showing the consequences of object-based anticipation in two similar, although distinguishable contexts.

5.6. Supplementary materials

Supplementary Table 1. ERP results.

Condition	ERP	t(sum)	p	D	Z	Electrodes
ATT NAME	P1	-46.98	<0.001	-0.54	14	P3' 'P7' 'O1' 'Oz' 'O2' 'P4' 'TP7' 'P1' 'P5' 'PO7' 'PO3' 'POz' 'PO4' 'PO8'
	P2	7.15	0.04		5	TP9' 'Oz' 'O2' 'TP10' 'PO8'
		-15.08	0.013		5	Pz' 'P3' 'P1' 'P5' 'P2'
	N2	-37.67	<0.001	-0.43	11	Pz' 'P3' 'P7' 'O1' 'P4' 'TP7' 'P1' 'P5' 'PO7' 'PO3' 'POz'
	P3	56.33	<0.001	0.51	18	Fp1' 'Fz' 'F3' 'F7' 'FT10' 'F4' 'F8' 'Fp2' 'AF7' 'AF3' 'AFz' 'F1' 'F5' 'FT7' 'FC4' 'F6' 'AF4' 'F2'
		-11.54	0.04		4	C3' 'CP1' 'C5' 'CP3'
ATT FACE	P2	-46.44	<0.001	-0.51	13	Pz' 'P3' 'P7' 'P4' 'P1' 'P5' 'PO7' 'PO3' 'POz' 'PO4' 'PO8' 'P6' 'P2'
	N2	-35.72	0.002	-0.42	13	P3' 'P7' 'P4' 'P8' 'P1' 'P5' 'PO7' 'PO3' 'POz' 'PO4' 'PO8' 'P6' 'P2'
	P3	16.9	0.017	0.43	8	C4' 'FC6' 'FC2' 'F4' 'FC4' 'FT8' 'F6' 'FCz'

Note: Only attention is shown, since expectation returned no significant differences. Results are reported for cued-uncued. Positive t values indicate larger amplitudes for cued stimuli, and negative indicate larger amplitudes for uncued stimuli. ATT = attention.

Supplementary Table 5.2. Time-frequency results

Condition	Freq.	Latency (ms)	t(sum)	p	D	Electrodes
Att_Face	Theta	200-500	-255.99	<0.001	-0.87	All
	Alpha	200-350	-54.75	<0.001	-0.59	'Fp1' 'Fz' 'F3' 'F7' 'FC6' 'F4' 'F8' 'Fp2' 'AF7' 'AF3' 'AFz' 'F1' 'F5' 'FT8' 'F6' 'AF8' 'AF4' 'F2'
		200-350	-32.17	0.003	-0.61	'CP1' 'Pz' 'P3' 'CP2' 'CP3' ‘P1' 'P2' 'CPz'
		400-650	-36.03	0.003	-0.65	'Pz' 'P3' 'P7' ‘O1' 'Oz' 'P1' 'P5' 'PO7' 'PO3' 'POz'
		400-650	-12.96	0.036	-0.4	'FC6' 'F8' 'FT8' 'F6' 'AF8'
Att Name	Theta	200-500	-197.01	<0.001	-0.702	'Fp1' 'Fz' 'F3' 'F7' 'FC1' 'T7' 'CP5' 'CP1' 'Pz' 'P3' 'P7' 'O1' 'Oz' 'O2' 'P4' 'P8' 'CP6' 'CP2' 'Cz' 'C4' 'T8' 'FC2' 'F4' 'Fp2' 'AF7' 'AF3' 'AFz' 'F1' 'F5' 'FT7' 'FC3' 'C1' 'C5' 'TP7' 'CP3' 'P1' 'P5' 'PO7' 'PO3' 'POz' 'PO4' 'PO8' 'P6' 'P2' ‘CPz' 'CP4' 'TP8' 'C6' 'C2' 'FC4' 'F6' 'AF8' 'AF4' 'F2' 'FCz'
	Alpha	400-650	-15.09	0.024	-0.44	'CP5' 'P3' 'CP3' 'P5' 'PO7' 'PO3'
	Gamma	300-650	60.05	<0.001	-0.45	'Fz' 'FC1' 'CP1' 'Pz' 'P3' 'P7' 'O1' 'Oz' ‘O2' 'P4' 'Cz' 'F1' 'C1' 'P5' 'PO7' 'PO3' 'POz' 'PO4' 'PO8' 'P2' 'CPz' 'FCz'
Exp Face	Theta	100-500	-125.43	<0.001	-0.58	'Fp1' 'Fz' 'F3' 'FC5' 'FC1' 'CP5' 'Pz' 'P3' 'P7' 'O1' 'Oz' 'O2' 'P4' 'P8' 'CP6' 'T8' 'FC6' 'FC2' 'F4' 'F8' 'AF7' 'AF3' 'AFz' 'F1' 'F5' ‘FC3' 'TP7' 'CP3' 'P1' 'P5' 'PO7' 'PO3' 'POz' 'PO4' 'PO8' 'P6' 'P2' 'TP8' 'FC4' 'FT8' 'F6' 'AF4' 'F2' 'FCz'
Exp Name	Theta	200-500	-8.55	0.047	-0.33	'Fz' 'FC1' 'F1' 'FCz'
	Gamma	300-650	23.89	<0.001	0.54	'Fz' 'F3' 'FC1' 'Cz' 'F1' 'FC3' 'C1' 'FCz'

Note: only significant clusters are shown. Reports of *t*-values and Cohen’s d are calculated for the time windows that where significant in a whole-brain time-frequency analysis

.

Supplementary Table 5.3. One sample *t*-test MD ROIs Univariate + Descriptives

ROI	Condition	Beta	SD	t	p
ACC	Cued_Exp	1.144	1.056	7.182	< .001*
	Uncued_Exp	1.281	1.357	6.260	< .001*
	Cued_Att	1.290	1.176	7.280	< .001*
	Uncued_Att	1.048	1.123	6.188	< .001*
AIFO	Cued_Exp	2.023	1.309	10.255	< .001*
	Uncued_Exp	1.971	1.688	7.746	< .001*
	Cued_Att	2.141	1.586	8.958	< .001*
	Uncued_Att	1.772	1.055	11.147	< .001*
IFS	Cued_Exp	1.369	1.387	6.547	< .001*
	Uncued_Exp	1.400	1.531	6.067	< .001*
	Cued_Att	1.305	1.340	6.463	< .001*
	Uncued_Att	1.365	1.012	8.943	< .001*
IPC	Cued_Exp	0.446	1.245	2.378	0.022*
	Uncued_Exp	0.593	1.424	2.760	0.008*
	Cued_Att	0.702	1.079	4.315	< .001*
	Uncued_Att	0.720	1.162	4.108	< .001*

Note: p-corrected for multiple comparisons = 0.025.

Supplementary Table 5.4. One sample *t*-test visual ROIs Univariate + Descriptives

ROI	Condition	Beta	SD	t	p
V1	Cued_Exp	3.771	2.111	11.848	<.001*
	Uncued_Exp	3.618	2.730	8.791	< .001*
	Cued_Att	3.318	1.860	11.834	< .001*
	Uncued_Att	3.127	1.987	10.439	< .001*
FG	Cued_Exp	3.536	1.488	15.759	< .001*
	Uncued_Exp	3.237	1.654	12.986	< .001*
	Cued_Att	3.467	1.494	15.394	< .001*
	Uncued_Att	2.867	1.302	14.604	< .001*
VVC	Cued_Exp	3.341	1.283	17.272	< .001*
	Uncued_Exp	3.126	1.649	12.575	< .001*
	Cued_Att	3.314	1.429	15.386	< .001*
	Uncued_Att	2.610	1.380	12.545	< .001*
LOC	Cued_Exp	2.401	1.595	9.984	< .001*
	Uncued_Exp	2.008	1.794	7.424	< .001*
	Cued_Att	2.080	1.581	8.727	< .001*
	Uncued_Att	1.781	1.499	7.880	< .001*

Supplementary Table 5.5. Post-hoc MD ROIs Univariate

ROI	**Block**	**T (cued-uncued)**	**P_{holm}**
ACC	expectation	-1.13	1.000
	attention	0.24	1.000
AI/FO	expectation	0.253	1.000
	attention	1.786	0.466
IFS	expectation	-0.174	1.000
	attention	-0.323	1.000
IPC	expectation	-0.730	1.000
	attention	-0.087	1.000

Note: t values are shown for cued – uncued comparisons in each ROI and condition.

Supplementary Table 5.6. Post-hoc visual ROIs Univariate

ROI	**Block**	**T (cued-uncued)**	**P_{holm}**
V1	expectation	0.497	1.000
	attention	0.663	1.000
FG	expectation	1.428	0.676
	attention	2.875	0.031
VVC	expectation	1.098	0.867
	attention	3.401	0.005*
LOC	expectation	1.962	0.253
	attention	1.520	0.545

Note: t values are shown for cued – uncued comparisons in each ROI and condition.

Supplementary Table 5.7. One sample *t*-test for decoding accuracies in MD regions + Descriptives

ROI	**Condition**	**Decoding accuracy**	**SD**	**t**	**p**
ACC	Cued_Exp	-4.261	24.403	-1.158	0.873
	Uncued_Exp	-1.705	19.363	-0.584	0.719
	Cued_Att	3.409	19.132	1.182	0.122
	Uncued_Att	-2.557	19.916	-0.852	0.800
AIFO	Cued_Exp	3.125	18.109	1.145	0.129
	Uncued_Exp	4.545	17.281	1.745	0.044
	Cued_Att	0.568	20.874	0.181	0.429
	Uncued_Att	-3.977	16.786	-1.572	0.938
IFS	Cued_Exp	8.807	20.278	2.881	0.003*
	Uncued_Exp	14.773	20.579	4.762	< .001*
	Cued_Att	8.807	22.648	2.579	0.007*
	Uncued_Att	1.989	21.557	0.612	0.272
IPC	Cued_Exp	4.261	20.521	1.377	0.175
	Uncued_Exp	13.068	20.345	4.261	< .001*
	Cued_Att	7.386	17.329	2.827	0.007*
	Uncued_Att	3.977	23.155	1.139	0.261

Note: *p* corrected for multiple comparisons = 0.012. Asterisks denote statistical significance. Decoding is reported as accuracy minus chance.

Supplementary Table 5.8. One sample *t*-test for decoding accuracies in visual regions + Descriptives

ROI	**Condition**	**Decoding accuracy**	**SD**	**t**	**p**
V1	Cued_Exp	7.955	20.009	2.637	0.012*
	Uncued_Exp	25.000	19.626	8.450	< .001*
	Cued_Att	19.318	19.710	6.501	< .001*
	Uncued_Att	17.045	20.369	5.551	< .001*
FG	Cued_Exp	24.148	20.424	7.843	< .001*
	Uncued_Exp	42.045	12.083	23.081	< .001*
	Cued_Att	37.500	14.517	17.134	< .001*
	Uncued_Att	34.091	17.127	13.203	< .001*
VVC	Cued_Exp	11.932	20.523	3.857	< .001*
	Uncued_Exp	29.261	20.521	9.459	< .001*
	Cued_Att	28.125	14.297	13.049	< .001*
	Uncued_Att	26.705	19.363	9.148	< .001*
LOC	Cued_Exp	11.364	17.433	4.324	< .001*
	Uncued_Exp	24.148	19.699	8.131	< .001*
	Cued_Att	21.307	17.796	7.942	< .001*
	Uncued_Att	18.466	16.060	7.627	< .001*

Note: *p* corrected for multiple comparisons = 0.012. Asterisks denote statistical significance. Decoding is reported as accuracy minus chance.

Supplementary Table 5.9. Post-hoc MD ROIs Decoding

ROI	Block	T (cued-uncued)	P_{holm}
ACC	expectation	-0.565	1.000
	attention	1.318	0.955
AI/FO	expectation	-0.356	1.000
	attention	1.141	1.000
IFS	expectation	-1.378	0.583
	attention	1.585	0.584
IPC	expectation	-1.923	0.336
	attention	0.744	1.000

Note: t values are shown for cued – uncued comparisons in each ROI and condition.

Supplementary Table 5.10. Post-hoc visual ROIs Decoding

ROI	Block	T (cued-uncued)	P_{holm}
V1	expectation	-4.669	0.013*
	attention	0.623	0.535
FG	expectation	-6.398	<0.001*
	attention	1.219	0.226
VVC	expectation	-4.966	<0.001*
	attention	0.407	1.000
LOC	expectation	-3.989	<0.001*
	attention	0.887	0.756

Note: t values are shown for cued – uncued comparisons in each ROI and condition.

Chapter VI
GENERAL DISCUSSION

The General Discussion is organized into several sections. First, we summarize the results obtained in the three studies of this thesis, and relate them to the goals and hypotheses described in Chapter 2. Then we consider alternative explanations for the results, followed by an integration of the results of the studies with the implications that they have for predictive coding and attention frameworks. Next, we present a general evaluation of the state of the art and the importance of reaching common terminology in cognitive psychology. We end with remaining questions and future directions.

6.1. General results summary

The aim of this thesis was to advance the knowledge of the nature of neural top-down preparation, by studying whether anticipatory and early perceptual representations are coded similarly in contexts of relevance and probability. We addressed this goal in three neuroimaging studies in which we collected EEG and fMRI data during pre- and post-stimulus intervals and used multivariate pattern analyses to evaluate the nature of represented information.

Our prerequisite goal **(G0)** was to design a task that allowed to study category anticipation in contexts of selection and probability. The task adequately fulfilled our objectives. It was able to elicit contexts of attention and expectation, as evidenced by the behavioral results. These (H1.1 and H2.1) were in line with previous paradigms that have measured neural activity during both attention and expectation, showing overall high accuracies (e.g. Scolari et al., 2012; Stokes et al., 2009) and improved performance for probable vs. improbable stimuli (e.g. Kok et al., 2012; Stein & Peelen, 2015). The two conditions were equated at perceptual, response and crucially, difficulty levels, evidenced by similar accuracies and reaction times across conditions. The task also allowed to apply multivariate analyses to study category representations during the anticipated period. The two stimulus categories, faces and names, elicited clear anticipatory and target-locked patterns of activity. Since all targets were different, these results were probably due to the representation of general categories, instead of specific stimuli. In addition, the use of different cues during the anticipation window allowed to control for perceptual confounds. Finally, this paradigm was robust across different neuroimaging methodologies and presentation times, as evidenced by the similar behavioral results obtained in the two studies.

In Chapter 3 we described the first study of this thesis, aimed **(G1)** at studying the time-resolved nature of brain anticipation by contrasting how information is represented in contexts of relevance and probability. In sum, our hypothesis **(H1)** that the neural codes in different conditions would be distinguishable was supported. Specifically **(H1.2)**, an RSA analysis revealed that during preparation several events take place. First, the cues' perceptual features are processed. Then, during most of the trial, global differences associated with the mechanisms deployed during attention and expectation take place. Finally, the anticipated category is preactivated with increasing sensitivity throughout the trial, in a ramping-up fashion. These results were replicated by time-resolved decoding, as both categories could be decoded during the anticipatory interval. Furthermore, when we controlled for perceptual biases by training and testing the classifier with different cues, we found results that were akin to those found for category in the RSA. Crucially, although both conditions exhibited similar decoding accuracies and ramping up effects, as well as large temporal generalization windows, **(H1.3)** cross-classification between them was not possible, a result that was not explained by the block design of the task. This indicates that although anticipated information is represented in the two conditions, the nature of such codes differs. Finally, **(G1.4)** we sought to explore the level of perceptual reinstatement. As expected **(H1.4)**, the two conditions differed in the pattern similarity between the cue and target, as evidenced by a CTT. The results showed that preparation during attention was more similar to target activity elicited by an independent localizer.

This first study left an important open question: what are the neural substrates underlying the representations found during anticipation? We approached this question in Chapter 4, where we describe a second study in which we adapted the behavioral paradigm to fMRI. Similar to Chapter 3, we first sought evidence of anticipatory representations. As we anticipated **(H2.2),** we observed anticipated coding that was different for attention and expectation. Univariate results revealed increased activations for attention in early visual regions, while the posterior cingulate cortex was more active for expectation. Moreover, decoding of the two blocks showed large differences in the recruitment of different visual and frontoparietal regions during anticipation, suggesting that different networks exhibit distinguishable mechanisms in these two contexts. Furthermore, in accordance to our prediction **(H2.3)**, the anticipated categories were represented in object-selective regions of the ventral visual cortex, partially replicating the results from the EEG study. We found that relevant and probable categories were represented in close, yet mostly different, brain regions. While the attended category was decoded mostly from ITG, probable categories were decoded in earlier regions, such as V1 or the lingual gyrus. Crucially, cross-classification between attention and expectation was again not possible, implicating that the coding nature of preparation across the two conditions is different **(H2.4)**. To explore the nature of these differences, we compared the level of similarity between anticipatory templates and targets. The cross-classification approach yielded results that partially matched **H2.5.** Although, as predicted, both attention and expectation generalized from preparation to target perception, they did so in different regions. Hence, we could not directly compare whether results were more pronounced for one of them. Moreover, a voxel selectivity ranking analysis yielded univariate evidence of the similarity of neural tuning during preparation and target perception. Contrary to our hypothesis, only expectation showed evidence of generalization from cue to target, and it did so in the VVC. Finally, we explored the behavioral relevance of neural representations during preparation. Contrary to our prediction **(H2.6)**, the distinctiveness of the prepared stimuli did not correlate with behavior in either condition. However, block decoding in frontoparietal regions showed a positive correlation with behavior, which might suggest that clear distinctions between expectation and attentional demands are associated with performance.

Finally, after studying the nature of top-down preparation, we aimed to study whether cueing sharpens or dampens sensory representations **(G3)**. We re-analyzed the data from the experiments in Chapters 3 and 4 to extract joint conclusions. The univariate cueing differences separately for attention and expectation yielded mixed evidence, whereas time-frequency analyses from EEG and GLM contrasts from fMRI showed converging evidence of increased activations for both probable and relevant stimuli. This, although in accordance with **H3.1** regarding attention, was opposed to the effects we expected to find for probable targets. Moreover, ERPs went in the opposite direction in attention, and were non-significant in expectation. To adjudicate between sharpening and dampening effects **(G3.2),** we compared the decoding accuracies for cued and uncued trials in attention and expectation using an ANOVA for both temporal generalization in EEG and specific ROIs in fMRI. Although evidence mostly supported a sharpening effect for attention (**H3.2)**, results did not reach significance in several contrasts. Crucially, probable targets were specially poorly decoded, compared to all other stimuli in both datasets, suggesting dampened representations for probable stimuli, even with larger activations. Additionally, we examined whether dampened representations would generalize better to the preparatory interval, since representations should be poorer in both cases. Our hypothesis **(H3.3)** was met, as expectation targets generalized to the cue, while relevant targets did not, again suggesting sharpening for attended and dampening for probable stimuli. In the fMRI study **(G3.4)** we studied the involvement of frontoparietal regions in sharpening and dampening. Although we could successfully decode target categories from IFS and IPC, these regions did not show any evidence of sharpening or dampening for either condition, contrary to **H3.4.** Finally, we examined the relationship between univariate and multivariate effects, and concluded that the results were not compatible with equivalent effects of similar effects for attention and expectation, suggesting that not only during anticipation, but also during target processing relevance and probability act through different mechanisms.

6.2. Top-down or bottom-up? Alternative explanations

Throughout this dissertation we have framed preparation as a top-down function: internal states induce preactivations that have consequences on perception and action (Brunia, 1999; Kveraga et al., 2007). In fact, we have shown that anticipatory representations can be decoded from temporal and occipital regions before target presentation (Study II). Moreover, these representations appear delayed in time, and increase towards the anticipated presentation time when it is previously fixed (Study I). However, could our results be a consequence of bottom-up, perceptually driven confounds and not top-down biasing?

Arguably, the possibility exists that these preactivations could be a consequence of bottom-up priming (Theeuwes, 2013). This alternative is based on studies that show that cueing a relevant location or object feature improves detection (Eger et al., 2007; Fernández et al., 2022; Theeuwes & Van Der Burg, 2007) by increasing cortical sensitivity. Following this proposal, results could be generated by the bottom-up perception of the visual cues, which would elicit neuronal firing in relevant perceptual regions. This activation would remain until the presentation of targets and would aid on its detection. However, automatic priming is particularly relevant in facilitatory effects of studies utilizing exact image cues representing the target object, in situations where the relevant or probable target dimension is consistently blocked, leading to its repetition across trials, or in exogenous spatial attention experiments (Battistoni et al., 2017). None of these features was present in our experiments. Our cues were symbolic shapes, different from anticipated targets. The predicted target category was randomized within each block, avoiding repetition effects that could last and be decoded from visual cortices. Moreover, cues predicted general categories (faces or names) but never individual stimuli, as the pool of target stimuli employed was quite large in numbers. Nevertheless, a priming account cannot be conclusively dismissed even when dealing with studies employing symbolic cues, such as words indicating the target object or previously learned associations between different stimuli in the same or different perceptual modalities (Feng et al., 2017). In such cases, processing the symbolic cue might activate the object's semantic representation, potentially resulting in an enhancement of perceptual processing for the primed object. However, automatic priming should lead to preactivations that are independent of specific task demands or top-down control. This is not compatible with research that shows sustained preparatory activity within the visual cortex is observable in working memory tasks with fixed delay periods, where information can be decoded with MVPA in a ramping-up manner, with increased accuracies as the target approaches (Serences et al., 2009; Stokes, 2015), a phenomenon labeled "activity silent working memory". Crucially, these findings appear only when information is actively maintained in WM, and decodable representations arise through top-down mechanisms, which fundamentally challenges the applicability of automatic priming accounts. This evidence resonates with our findings in Study I that decoding only emerges as the target approaches, which is difficult to explain with mere automatic priming without relying on top-down mechanisms.

Decoding evidence of sustained representations in the absence of perceptual stimulation has also been challenged. Specifically, it has been suggested that systematic eye movements, or micro-saccades, during the delay period could contaminate the decoding results (Mostert et al., 2018; Quax et al., 2019; Thielen et al., 2019). Arguments in favor of this explanation come from data that shows that decoding results based on neural information are similar to the results based on eye-tracker data, which is a prominent concern in M/EEG studies. There are three main mechanistic explanations as to how eye movements could bias decoding results (Mostert et al., 2018). First, the rotation of the eyeball is picked by EEG sensors, due to its positively charged cornea and negatively charged retina, which together act as a dipole. This artifact is only a concern in electrophysiology studies, since it cannot be detected by fMRI. Its effects can be reduced by removing that part of the signal by performing independent component analysis (ICA), or regressing out activity explained by an electrooculogram (EOG) or, better, eye-tracker data (Thielen et al., 2019). In Study I, we used ICA to remove eye-movement components. Moreover, an electrophysiological bias would drive classification weights to frontal electrodes, which is not our case (Figure 3.3, evidence of the topographical distribution of RSA category and cue shape results). Second, saccades biased to a specific part of the screen imply that the perceived image could change. For instance, if there is a fixation point, which is the norm in psychophysics experiments, classifiers could be decoding the relative position of the fixation point in the visual field. This confound would also affect fMRI data, and would lead to decoding in visual regions or weights in occipital sites, closer to our results. However, tasks where eye movements are unlikely greatly reduce this effect (Mostert et al., 2018). This includes presenting all stimuli in the same position, manipulating features that are less directional (such as grating orientations), or showing targets in a random position that cannot be predicted by the cue. Again, our results can hardly be explained by this effect, since all targets were presented at center of the screen. Still, it could be argued that faces and names can elicit different types of eye movements. Faces could drive saccades to the approximate position where the eyes will appear (above the fixation) while names could elicit lateralized movements. Although future revisions of the task where eye tracker data are collected are needed, this phenomenon alone still cannot explain our results. First, fMRI showed decoding in object-selective rather than early visual regions. Second, perceptually driven bottom-up effects should be similar across different cognitive contexts, which does not agree with cross-decoding results. And third, several analyses showed pattern similarity between anticipation and perception in both EEG and fMRI, which reveal that anticipation generated partially similar templates to the anticipated category in visual regions. Finally, eye-movements would also elicit activity changes in regions related to eye position, such as the FEF. In fMRI experiments, significant decoding in this region should rise concerns and prompt other control analyses. However, we did not find significant coding in that region during whole brain analyses. Altogether, although research aiming at studying top-down anticipations should make efforts to control for these possible effects and account for systematic biases, our evidence speaks strongly against the presence of these biases in our results.

6.3. Preparation as a heterogenous state

In Studies I and II we consistently showed that although anticipated information is coded in the brain, the exact nature of said coding differs. This adds to the literature that has studied a plethora of mechanisms taking place in the absence of perceptual stimulation, such as imagery (Cichy et al., 2012; Dijkstra et al., 2019), working memory (Nir-Cohen et al., 2020; Stokes, 2015), cognitive control (González-García et al., 2021; Palenciano, González-García, Arco, Pessoa, et al., 2019), selective attention (Battistoni et al., 2017; Woolgar et al., 2015) or probabilistic expectations (de Lange et al., 2018; Kok et al., 2017).

All these processes share two characteristics: they are instantiated through top-down endogenous mechanisms, and they induce the representation in perceptual brain regions of stimuli that are not present. The extent to which they consist of, in fact, a single unitary function or different manifestations of several mechanisms is debated. Interestingly, the research trend with attention and WM, which have been extensively compared (see Gazzaley & Nobre, 2012 for a review), is somewhat the opposite to our attention-expectation comparison. In that case, attention and WM have been usually regarded and studied as different phenomena and, only relatively recently, research has started to show that their neural bases are similar, and actually selection can account for a large part of the typical WM effects during the delay period, such as capacity, cortical representations and FPN involvement (Chun, 2011; Postle, 2006). In fact, it has been proposed that working memory is better understood as the result of the interaction between cognitive control, attention and perceptual representations (Bledowski et al., 2010). Working memory would not be a process, but a consequence: cognitive control displays attention, which in turn induces the representation of relevant stimuli in visual regions, thus “creating” working memory.

Following this rationale, it could be argued that preparatory attention and expectation would be similar manifestations of top-down regulations (Gazzaley & Nobre, 2012). In fact, their neural bases have been shown to be partially overlapping: both processes rely on FPN or MDN regions (Esterman & Yantis, 2010; Rahnev et al., 2011). In attention, they would be the sources of top-down templates (Desimone & Duncan, 1995), that are represented in sensory regions. In expectation, they would be an intermediate region, while anticipatory templates would be generated in memory sources such as the hippocampus (Barron et al., 2020; Hindy et al., 2016; Kok & Turk-Browne, 2018). Our data similarly showed that category anticipation accuracies were similar (see Figure 3.4) in both conditions, although we did not find evidence of frontoparietal involvement. However, in Study I we showed that preparation is a complex phenomenon that entails several mechanisms, from cue processing to category anticipation. Moreover, the patterns that support anticipatory representations are different across contexts.

Furthermore, the anticipatory relevance vs. probability showed different similarity between preparatory representations and target perception, although the results across different methodologies where not consistent: whereas the CTP in EEG showed more reinstatement for attention; the slope analysis in fMRI revealed more similar tuning for expectation in VVC; and cross-classification showed more involvement for expectation on EEG, and different, barely overlapping regions in fMRI. Although the CTP was the only analysis that showed larger reinstatement for attention, it was also the only one that was carried out using an independent localizer. In all other analyses, perceptual data came from target epochs. That is, it is possible that anticipated information during expectation is better represented, as it is during target perception, or conversely, as shown in Chapter 5, it could be that representations during probabilistic contexts in target perception are attenuated and thus more similar to anticipation. So, taking only CTP results from EEG as the most reliable source of information, the conclusion is still that anticipatory templates are different across attention and expectation, and that differences could not be explained by perceptual or difficulty differences across the two conditions.

Akin to what has been suggested for WM and attention, our results collectively indicate that preparation is not a specific function, but rather a temporal window where several neural mechanisms unfold. This includes a fundamental tool: top-down templates lead to preactivations in sensory cortices. However, the nature of these preactivations depends on the context. If anticipation happens in a context of perceptual expectations, mechanisms pertaining to expectations are deployed. If it happens in a context of relevance, attentional resources can bias sensory cortices towards attended stimuli. Moreover, it would be logical to think that both attention and expectation can interact in that anticipatory space, as it has been repeatedly been suggested by studies that have examined their combined effects during target processing (e.g. Jiang et al., 2013; Richter & de Lange, 2019). Research that examines their interaction during the preparatory interval is still lacking.

6.4. General categories vs. specific stimuli

In all three experimental chapters we attempted to examine the relationship between anticipation and target perception. Our first focus was to examine whether anticipatory templates were similar to perceived images. Crucially, our manipulation was oriented to induce anticipations for broad categories, not individual stimuli. Participants were unaware of the gender or identity of the faces and words before they were perceived. Previous studies of category anticipation or imagery have employed a much narrower number of images. For instance, Esterman & Yantis, (2010) compared category expectations to faces and houses, but only used 12 different stimuli of each. Cichy et al. (2012) in an imagery task with four different categories used three individual images per condition. Lee et al. (2013) also compared the representation of different categories in a WM task, but used two stimuli per category. In our task, participants saw a different target on every trial, so it was not possible for them to anticipate the exact features of the incoming stimulus. Thus, here we show that templates associated with whole perceptual categories can also be represented in the brain. This might also contribute to our overall low decoding accuracies, since anticipatory templates where probably more abstract than those in similar previous experiments with more repetitions of each stimulus (e.g. Cichy et al., 2012; Esterman & Yantis, 2010). Crucially, these general representations took place in category-selective cortices, and shared activity patterns with the more detailed visual targets. Nevertheless, this manipulation had another consequence: in all trials there was a possible component of surprise, as the exact stimulus could not be predicted. Even so, in Chapter 5 we showed how dampening effects in expectation are robust and take place even when the prior pertains to general categories. It has been proposed that for an effect of surprise that could deploy attentional resources and possibly reverse the dampened representations we found here, the level of unexpectedness needs to be large enough to cross a perceptual salience threshold (Press et al., 2020), which was hardly the case in our task. Altogether, our results suggest that anticipatory templates induce changes in neural tuning depending on the cognitive context. Crucially, this template is not necessarily equal to individual concrete stimuli. It induces changes in visual processing even when the exact nature of the stimulus is not known.

6.5. How does preparation relate to perception?

Although we showed that information is at least similarly represented during anticipation and target perception, our results do not allow for a straightforward interpretation of the spatiotemporal dynamics and changes that take place between the two stages. In Chapter 4, we could not link category anticipations with behavioral performance, unlike previous reports (Auksztulewicz et al., 2017; González-García et al., 2017; Jin et al., 2020).

Moreover, although sensory templates generalized from cue to target, decoding of context (attention vs. expectation) did not. Hence, we cannot conclude that the patterns of brain activity that support mechanisms for anticipatory attention and expectation other than stimulus representations are similar to the ones that arise during target perception and interpretation. Similar results where anticipatory and processing activity are not fully equivalent have been found by Panichello & Buschman, (2021), which found that preparatory attention and visual attention generalized (cross-classified) in PFC, but not in V4, parietal cortex or FEF. This highlights the complexity of visual processing, and how different mechanisms are involved from anticipation to action depending on several events. Although more research is needed, below we speculate about what processes may unfold differently in attention and expectation during a complete trial.

As revealed by RSA in Chapter 3, when the cue is presented, its perceptual features are processed, and then mechanisms pertaining to either condition are deployed. In attention trials, cognitive control would then guide attention to relevant features (Diamond, 2013). The connection between the cue and cognitive control mechanisms could be exerted through associative learning after several repetitions of cue-target-feedback pairs during the practice blocks. The role of associative learning in cognitive control is possibly to induce the activation of several mechanisms which allow flexible behavior, including attention orienting (Abrahamse et al., 2016; Braem & Egner, 2018). The orienting of attention to different stimulus categories (Aranda et al. 2010) is then implemented via top-down templates which would be generated in FPN regions, such as IFS, IPS, or PFC (Battistoni et al., 2017; Kim & Kastner, 2019; Panichello & Buschman, 2021) and are represented afterwards in category selective areas (Peñalver et al., 2023; Chapter 4). If it is known when visual stimulation will happen, the representations could be maintained in FPN areas, and be intensified in downstream regions as target approaches and a template to compare is more necessary. According to predictive coding and biased competition models, (Feldman & Friston, 2010; Kok, Rahnev, Jehee, Lau, & de Lange, 2012), attention would deploy these representations in superficial cortical layers (Muckli et al., 2015), by increasing neural excitability and through, perhaps, top-down inhibition of neurons tuned away from the relevant category. This might explain the increased activity found for attention in visual regions (Figure 4.5). Once the target appears, if it matches the anticipated category, neural precision (tuning to the relevant category) is increased (Feldman & Friston, 2010), thus increasing the weight of relevant sensory information, inducing sharpened representations (Woolgar et al., 2015). However, we could not find convincing evidence on this last direction. This may be due to all targets capturing at least some attention due to their novelty and the absence of a distractor stimulus to drive selection away.

In expectation, rather than goal directed cognitive control, cues could trigger automatic associations, learned through experience (practice) about the probable characteristics of incoming stimuli (Kok et al., 2017). This probabilistic knowledge could be stored in the hippocampus (Aitken & Kok, 2021; de Lange et al., 2018), which could then generate perceptual templates in visual regions. Again, these templates would be active in superficial layers of the cortex. Expectation would be more automatic, not requiring specific control. Moreover, preactivations would be maintained in earlier (Chapter 4), less stimulus specific regions, and without a pre-inhibition of neurons tuned to other stimulus categories. When the target is presented, if it matches the preactivated template, similar characteristics would be filtered out, and representations would be dampened to reduce computational costs (Chapter 5), which could help to give faster and more accurate responses. This filter would make probable stimuli less specific, and hence more similar to category anticipation in earlier sensory regions. Importantly, the ability of unexpected stimuli to elicit an attention-like response and induce a sharpened representation would depend on the level of surprise and learning value (Press et al., 2020), which in our task was possibly not enough, making the representations similar between unexpected and both attended and unattended stimuli.

6.6. Who knows what attention is?

This dissertation results from our effort to expand and detail the definitions of key concepts in cognitive science: preparation, attention, and expectation. In this last section, we would like to divert a bit from specific theoretical models and results and make a very brief, more general evaluation of the state of the art in relation with our work.

As stated at the introduction, the quote “Everyone knows what attention is” (James, 1890) may not have aged too well. Perhaps this is because in 130 years of cognitive research there are still not enough different and consistent labels that describe the diverse set of empirical results that are stored under the umbrella of “attention”. As discussed by Hommel (2019), who counterargued that “no one knows what attention is”, attention has been fractioned to include a plethora of processes related to a cognitive system with limited capacity that needs to focus resources on *relevant* information. This includes the selection of external events to be processed, inhibition of distractors, focus on salient information, increased representations of locations or features, among many others. It is, as claimed by the authors, very unlikely that all these phenomena share the same functional mechanisms. Moreover, research is grounded on concepts that may not be systematically or properly defined. Such could be the case of “relevance”. When experimenters manipulate probability to measure attention (e.g. Posner, 1980), they are assuming that something probable is also relevant: knowing where a stimulus can appear drives attention to that location. But one could logically argue in the opposite direction: something unlikely is more relevant, which is why, for instance, oddball paradigms can be used to capture attention (Alink & Blank, 2021; Feuerriegel, Yook, et al., 2021). Both definitions can be logically and arguably correct, which has made the joint interpretation of both types of results difficult (Press et al., 2020).

Moreover, this lack of a systematic definition of the effects associated with attention has potentially led to split lines of research that grow increasingly harder to re-join as literature expands. Going back to the topic of this thesis, the idea that selection and probability may be different cognitive processes gained prominence when Summerfield & Egner (2009) highlighted that both concepts had been studied as one unitary process, when they were not necessarily equivalent. Although there is an extensive body of research that has compared how attention and expectation interact and differ (e.g. Auksztulewicz et al., 2017; Jiang et al., 2013; Wyart et al., 2012), and additional new evidence suggestive of a different set of mechanisms (Peñalver et al., 2023), there is still an active line of research that continues manipulating probability to measure attention (e.g. Battistoni et al., 2017). Both lines keep advancing mostly in parallel, studying two potentially different processes and labeling them in the same manner. Thus, when designing new tasks and discussing new sets of results, the inferences can be not only biased in opposite directions depending on the literature review, but are possibly weakened by conflicting results, which complicate the integration of information.

This issue has been discussed several times and can be applied to other cognitive domains such as memory or consciousness (Hommel et al., 2019). However, there could be another characteristic of cognitive science that underlies this circumstance: the terms that we use to describe brain/cognitive function have not really changed, or not as much as in other scientific disciplines, since William James (1890). The work of Poldrack (2016) has shown that the terminology that we use today matches around 20% of the one that appeared in James' Principles of Psychology. Comparatively, only 0.09% of the terms used in Biology matched the ones used in similarly influential textbooks of the same period. Other reports have found that in specific contexts the share of prominent concepts in today's literature with James can be to up to 47% (Ortiz-Tudela & González-García, 2023). Hence, if terminology is not updated, more confusion is bound to happen, which inevitably hinders scientifical advances. These authors propose, as a tentative explanation, that in psychology there has not been any major breakthrough that could justify the opening of new lines of research and the use of updated terminologies.

Interestingly, possibly one of the most similar things to a major change in cognitive neuroscience has been the importance of predictive processing (Feldman & Friston, 2010; Friston, 2005). Indeed, this thesis is conceptualized within this integrative framework, given its flexibility and compatibility with attentional theories such as the biased competition (Desimone & Duncan, 1995). This has allowed us to integrate both attention and expectation within the same framework and attribute similar mechanistic explanations to two different sets of results. However, there is some fair criticism (Litwin & Miłkowski, 2020) to the speed to which neuroscience has attempted to fit much of the effects we observe within a framework that is still in development. Examples of this are common (or not so common) effects such as expectation suppression, which may not be understood enough to guide such a wide and still conflicting (e.g. dampening vs. sharpening interpretations) set of results (Feuerriegel et al., 2021; Press et al., 2020).

Rather than claiming that psychological science, predictive processing theories, or the term attention are not crucial anymore, we argue that a more systematic way to define terms (including concepts, effects or tasks; Poldrack et al., 2011) would likely help the advance of research. This is, a more data-driven approach that clearly associates tasks (not necessarily concepts) with results could eventually lead to an easier understanding of behavior and brain function, and the definition of more precise integrative theories.

6.7. Remaining questions and future work

Despite the contributions of this thesis, there are still several questions that remain unanswered. First, although we systematically compared the patterns that support category representation in attention and expectation, and argue that its perceptual nature might explain the different patterns of activity that we found, the exact mechanisms supporting these differences is still unknown.

Our evidence was conflicting regarding the similarity between preparation and target processing. The inclusion of a localizer block in the fMRI experiment could help painting a more precise picture in this direction.

Different oscillatory patterns could be supporting each type of representation. It has been suggested that top-down influences are implemented in external cortical layers through slower oscillations, while gamma might support bottom-up representations (Fries, 2005; Fries et al., 2001). This could be examined via MVPA analyses which, instead of voltage, use power of different oscillatory bands as features to train the classifier (see Jafarpour et al., 2013 for a similar approach). Conversely, specific frequencies could be filtered out from the EEG signal and the analyses repeated to examine which oscillations contribute most to decoding results (López-García et al., 2020, 2021).

Moreover, the contribution of top-down regions to the effects showed here is unclear. In the fMRI study, the MD ROIs included in Chapter 5 (IFS, ACC, IPC, AI/FO) could be defined a priori, plus the hippocampus, given its suggested relevance in expectation contexts (Aitken & Kok, 2022; de Lange et al., 2018; Kok et al., 2017). Furthermore, approaches such as Dynamic Causal Modelling (Auksztulewicz & Friston, 2015) have been employed within predictive coding to study the influence of higher-order regions on neural activity in downstream areas, and vice-versa. Particularly, we could study the specific dynamics that allow for the implementation of specific anticipatory templates in sensory regions, and the probable neural mechanisms that support them.

The experiments of this thesis manipulated attention and expectation in the same way, using the same stimuli but exploiting the capabilities of each neuroimaging technique. Although we compared the results of both in Chapter 5, it is possible to go one step forward. Fusion RSA (Cichy & Oliva, 2020; Hebart et al., 2018) allows to directly examine the representational similarity of two different measures, such as EEG and fMRI. This would allow us to obtain a more detailed spatial and temporal pattern of brain activity during the preparatory and perceptual interval. At the moment of writing of this dissertation, this analysis is still in preliminary stages, but shows promising results that could complement the studies already presented (Figure 6.1). Further developments could even apply the same rationale of using the power of specific frequency bands described for MVPA, but for fusion analyses, laying out how oscillations are related to anticipatory processing in different regions.

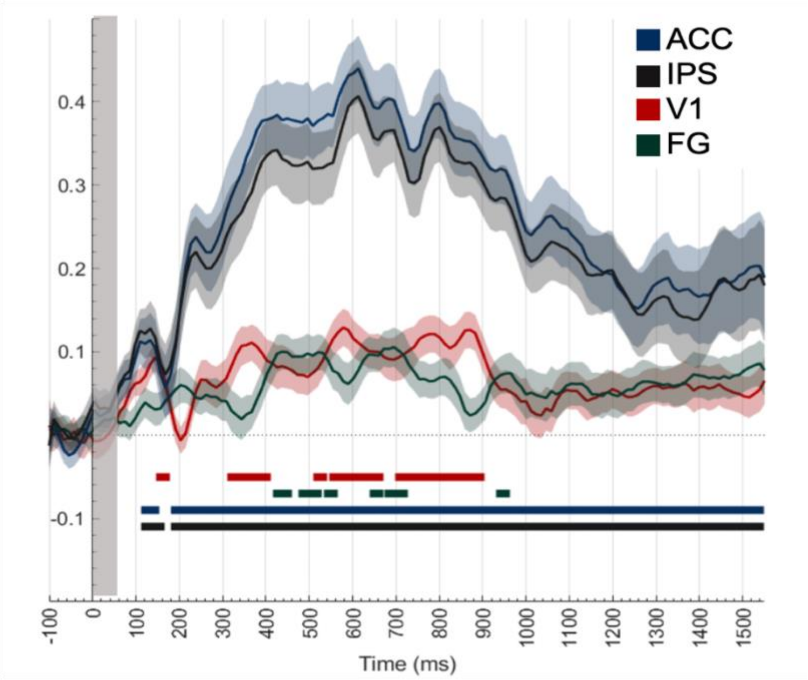


Figure 6.1. Preliminary Fusion-RSA results for the preparation window. Each line corresponds with a ROI. Shading indicates the standard deviation. Grey column shows the cue's onscreen time. Horizontal colored lines show statistical significance after a cluster-based permutation analysis. The regions shown in this figure are: aCC = anterior cingulate cortex, IPS = inferior parietal sulcus, V1 = primary visual cortex, FG = fusiform gyrus.

Finally, we have shown and discussed that several of the effects found are highly dependent of the specific task that is used to measure them (Feuerriegel, Vogels, et al., 2021; Richter & de Lange, 2019). Future designs could change characteristics of the task to explore the robustness of these findings. For instance, reducing the number of stimuli could increase the effects of the anticipatory representations found here. Moreover, to examine how we can flexibly move from a selection to a probability context, both conditions could be manipulated within the same blocks. This would help understanding what are the mechanisms that allow, not only for different representations in attention and expectation, but also the change from one context to the other.

In summary, with this thesis, we have provided new evidence and insight into the complex phenomenon that is top-down preparation. We have shown that it entails a plethora of diverse mechanisms, including the representation of anticipated perceptual categories in specific regions. Crucially, we have evidenced how the nature of these representations can change when the cognitive contexts are different, such as those involving selective attention and expectation. Future research will face the challenges of further advancing our knowledge of the mechanisms that support neural anticipation and understanding how they contribute to adaptive and efficient behavior.

Resumen

El mundo es complejo y está en constante cambio. Somos bombardeados por un gran número de estímulos que generan patrones ambiguos en nuestros receptores sensoriales, a los que normalmente necesitamos dar respuestas rápidas y eficaces. Para hacer frente a esta complejidad y ambigüedad, entre otras cosas los humanos utilizamos la capacidad de anticipar lo que va a ocurrir antes de que ocurra. De hecho, está muy extendida la idea de que prepararse para percibir información inminente mejora el rendimiento (Barton et al., 2006). Sin embargo, de estas afirmaciones aparentemente evidentes surgen algunas preguntas fundamentales. ¿Cómo mejora la preparación el rendimiento? ¿Implica la representación de información específica? Y, sobre todo, ¿es la preparación proactiva un proceso unificado o emplea mecanismos disociables en diferentes contextos cognitivos?

Esta tesis enmarca la preparación como un proceso cerebral de origen interno o endógeno, de arriba a abajo (en adelante, *top-down*), en referencia al origen de la información en regiones superiores de la corteza, que descenderían áreas perceptivas. Esta definición contrasta con otra rama de mecanismos de abajo a arriba (*bottom-up*), en los que el desencadenante de la actividad está en los estímulos del entorno. Hay al menos dos posibles fuentes de información que pueden llevar a esta anticipación endógena: nuestras metas y objetivos, y el conocimiento que tenemos de las regularidades del entorno.

La actividad asociada a metas se relaciona el control interno de recursos específicos que nos permiten un desempeño eficaz de diferentes tareas. Este control cognitivo proactivo (Braver, 2012) se asocia a regiones cerebrales como la red frontoparietal o la red de múltiple demanda (Dosenbach et al., 2007). Entre sus principales capacidades encontramos la orientación de la atención hacia información relevante. Esta atención selectiva actúa como un filtro que nos permite procesar mejor la información necesaria mientras ignoramos aquella que nos distrae. Crucialmente, se ha comprobado en numerosas ocasiones que mientras seleccionamos información relevante, dicha información se representa en la corteza visual, a través de patrones de actividad asociados a los estímulos a los que estamos atendiendo (Peelen &

Kastner, 2011; Stokes et al., 2009). Más aún, cuando se presentan estímulos que se corresponden con lo seleccionado, se suelen encontrar tanto aumentos de actividad en regiones perceptivas (Kastner et al., 1999) como mayor precisión en el procesamiento de dichos estímulos (Woolgar et al., 2015).

Por otra parte, los modelos que explican los efectos del conocimiento previo sobre las regularidades estadísticas del entorno, o expectativas, han ganado mucha relevancia en los últimos años. Estos modelos de procesamiento predictivo (Feldman & Friston, 2010; Friston, 2005) destacan cómo en todo momento el cerebro actúa como una máquina que anticipa lo que es más probable. Estas predicciones *top-down* se contrastarían con la información de los receptores sensoriales, *bottom-up*. Del contraste de ambas surgen errores de predicción, que se utilizan para actualizar las siguientes predicciones y así dar al cerebro un mecanismo flexible y capaz de adaptarse en diferentes circunstancias. Estos modelos también describen cómo la información probable puede representarse en la corteza visual antes de que se de estimulación *bottom-up* (de Lange et al., 2018; Kok et al., 2017). Sin embargo, al contrario que en atención selectiva, aquí cuando se percibe un estímulo probable se suelen encontrar menores activaciones (Feuerriegel, Vogels, et al., 2021). Con respecto a la precisión de estas representaciones, la evidencia es aún conflictiva, y se han encontrado datos que apuntan tanto a una mayor precisión (Kok et al., 2012) como a un empobrecimiento de dichas representaciones (Richter et al., 2018).

Aunque estos dos fenómenos, atención selectiva y expectativa, se asocian a efectos diferentes durante el procesamiento de estímulos, en ambos se dan representaciones anticipatorias en la corteza visual. Sin embargo, no se ha investigado si estas representaciones comparten los mismos mecanismos de procesamiento, o si por el contrario la preparación es un fenómeno complejo y heterogéneo que depende del contexto cognitivo en el que se produzca.

Por ello, el objetivo principal de esta tesis es examinar la naturaleza de la preparación, y evaluar cómo cambia a través de diferentes contextos. Para ello diseñamos una tarea en la que diferentes señales daban información sobre estímulos objetivo. Dichas señales podían indicar la relevancia (atención) o probabilidad (expectativa) de los estímulos. Esta tarea se implementó en dos experimentos en los que se recogieron datos de distintas herramientas de neuroimagen. Los análisis se centraron fundamentalmente en técnicas multivariadas (Hebart & Baker, 2018), particularmente sensibles y eficaces en el estudio de patrones de actividad cerebral que informan sobre cómo se representa la información.

En los dos primeros estudios, investigamos la preparación proactiva, explorando cómo se representa en el cerebro la información anticipada, y si cambia, y cómo, a través de contextos de relevancia y probabilidad. El primer estudio se centró en la dinámica temporal de la actividad preparatoria. Para ello la actividad cerebral se registró mediante Electroencefalografía (EEG). Los resultados mostraron que durante la ventana preparatoria tienen lugar varios acontecimientos. En primer lugar, se procesan las características perceptivas de la señal. Después, durante la mayor parte del ensayo se encontraron diferencias globales asociadas a los mecanismos desplegados durante la atención y la expectativa. Por último, la categoría anticipada se representó a lo largo del ensayo, especialmente al final. De forma crucial, aunque tanto atención como expectativa mostraron precisiones de decodificación y efectos de aumento similares, los resultados de una no generalizaron a los de la otra. Esto sugiere que, aunque la información anticipada se representa en las dos condiciones, la naturaleza de tales representaciones difiere. Por último, tratamos de explorar el nivel de similitud entre las representaciones preparatorias y la percepción de los estímulos. Las dos condiciones diferían en la similitud de patrones entre la señal y el objetivo, concretamente, la preparación durante atención era más similar a la actividad del objetivo.

Este primer experimento dejó una importante cuestión abierta: ¿cuáles son los sustratos neurales que subyacen a las representaciones encontradas durante la anticipación? Abordamos esta cuestión en el segundo experimento, en el que adaptamos el paradigma conductual a la resonancia magnética funcional (RMf). De nuevo, encontramos pruebas de codificación anticipada que eran diferentes para atención y expectativa. Los resultados revelaron un aumento de las activaciones para atención en las regiones visuales tempranas, mientras que el córtex cingulado posterior estuvo más activo para la expectativa en esta ventana. Además, la información anticipada mostró grandes diferencias en la implicación de diferentes regiones visuales y frontoparietales durante la anticipación, lo que sugiere que diferentes redes exhiben mecanismos distinguibles en la atención y la expectativa. Además, encontramos datos indicativos de la representación en la corteza visual ventral, replicando en parte los resultados del experimento EEG. Descubrimos que la categoría atendida y la probable estaban representadas en regiones cercanas, aunque en su mayoría diferentes. De nuevo, la clasificación cruzada no fue posible entre atención y expectativa, lo que implica que la naturaleza representacional de la preparación en las dos condiciones era diferente. Al comparar la similitud entre las representaciones anticipatorias y la percepción de los estímulos, encontramos que estas similitudes se dieron en regiones diferentes según la condición. Además, un análisis de selectividad de vóxeles, similar al de Richter et al. (2018), reveló que solo expectativa mostraba alguna evidencia de generalización de cue a target, y lo hacía en el la corteza visual ventral. Por último, la representación de los estímulos anticipados no se correlacionó con el comportamiento en ninguna de las condiciones.

Por último, tras estudiar la naturaleza de la preparación *top-down*, investigamos cómo esta cambiaba la precisión de las representaciones inducidas previamente a través de la atención y la expectativa. Específicamente, nos propusimos estudiar si las señales agudizan o amortiguan las representaciones sensoriales. Para ello, volvimos a analizar los datos de los dos experimentos para extraer conclusiones conjuntas. Como paso preliminar, nos propusimos encontrar pruebas de diferencias de activación por separado para la atención y la expectativa. Estos análisis mostraron pruebas convergentes de un aumento de las activaciones tanto para los estímulos probables como para los relevantes. Respecto a los efectos de agudización y amortiguación, los resultados apoyaron mayoritariamente un efecto de agudización de la atención, y una marcada amortiguación en expectativa. Además, examinamos si las representaciones amortiguadas se generalizarían mejor al intervalo preparatorio, ya que las representaciones deberían ser más pobres en ambos casos. Nuestra hipótesis se cumplió, ya que los ensayos de expectativas generalizaron a la señal, mientras que los objetivos relevantes no lo hicieron. Esto sugiere de nuevo una agudización para los estímulos atendidos y una amortiguación para los probables. En el experimento fMRI estudiamos la implicación de las regiones frontoparietales en los efectos de agudización y amortiguación. Aunque pudimos decodificar con éxito las categorías objetivo en algunas zonas frontoparietales, estas regiones no mostraron ninguna evidencia de agudización o amortiguación para ninguna de las dos condiciones. Por último, concluimos que los resultados no eran compatibles con efectos equivalentes para la atención y la expectativa, lo que sugiere que no sólo durante la anticipación, sino también durante el procesamiento del objetivo la relevancia y la probabilidad parecen actuar a través de mecanismos diferentes.

En conjunto, nuestros resultados muestran que la preparación es un evento complejo, en el que la representación de diferentes estímulos depende en gran medida del contexto en el que se den. Esta preparación, además, afecta a la fidelidad con la que se representan los estímulos percibidos, lo que también depende del contexto. Finalmente, esta tesis resalta las diferencias en la atención selectiva y la expectativa, dos procesos que, aunque comparten similitudes, conllevan mecanismos y representaciones diferentes.

Abstract

The world is complex and constantly changing. We are bombarded by a large number of stimuli that generate ambiguous patterns in our sensory receptors, to which we usually need to give quick and effective responses. To cope with this complexity and ambiguity, among other things, humans use the ability to anticipate what is going to happen before it happens. Indeed, it is widely believed that preparing to perceive impending information improves performance (Barton et al., 2006). However, some fundamental questions arise from these seemingly self-evident claims. How does preparation improve performance? Does it involve the representation of specific information? And, most importantly, is proactive preparation a unified process or does it employ dissociable mechanisms in different cognitive contexts?

This thesis frames preparation as a brain process of internal or endogenous origin, top-down, referring to the origin of information in higher regions of the cortex, which would descend perceptual areas. This definition contrasts with another branch of bottom-up mechanisms, in which the trigger of the activity is in the stimuli of the environment. There are at least two possible sources of information that can lead to this endogenous anticipation: our goals and objectives, and the knowledge we have of the regularities of the environment.

The activity associated with goals is related to the internal control of specific resources that allow us to effectively perform different tasks. This proactive cognitive control (Braver, 2012) is associated with brain regions such as the frontoparietal network or the multiple demand network (Dosenbach et al., 2007). Among its main capabilities we find the orientation of attention towards relevant information. This selective attention acts as a filter that allows us to better process necessary information while ignoring distracting information. Crucially, it has been shown on numerous occasions that while we select relevant information, this information is represented in the visual cortex, through patterns of activity associated with the stimuli we are attending to (Peelen & Kastner, 2011; Stokes et al., 2009). Moreover, when presented with stimuli that match what is selected, both increases in activity in perceptual regions (Kastner et al., 1999) and greater accuracy in processing those stimuli are often found (Woolgar et al., 2015).

On the other hand, models that explain the effects of prior knowledge on the statistical regularities of the environment, or expectations, have gained much relevance in recent years. These predictive processing models (Feldman & Friston, 2010; Friston, 2005) highlight how at all times the brain acts as a machine that anticipates what is most likely. These top-down predictions would be contrasted with bottom-up information from sensory receptors. From the contrast of the two, prediction errors emerge, which are then used to update the following predictions and thus give the brain a flexible mechanism capable of adapting in different circumstances. These models also describe how probable information can be represented in the visual cortex before bottom-up stimulation (de Lange et al., 2018; Kok et al., 2017). However, in contrast to selective attention, here when a probable stimulus is perceived, smaller activations are usually found (Feuerriegel et al., 2021). With respect to the precision of these representations, the evidence is still conflicting, and data have been found pointing to both sharpened accuracy (Kok et al., 2012) and a dampening of these representations (Richter et al., 2018).

Although these two phenomena, selective attention and expectation, are associated with different effects during stimulus processing, both involve anticipatory representations in visual cortex. However, it has not been investigated whether these representations share the same processing mechanisms, or whether, on the contrary, preparation is a complex and heterogeneous phenomenon that depends on the cognitive context in which it occurs.

Therefore, the main goal of this thesis is to examine the nature of preparation, and to evaluate how it changes across different contexts. To this end, we designed a task in which different cues provided information about target stimuli. These cues could indicate the relevance (attention) or probability (expectation) of the stimuli. This task was implemented in two experiments in which data were collected from different neuroimaging tools. The analyses focused primarily on multivariate techniques (Hebart & Baker, 2018), which are particularly sensitive and effective in studying patterns of brain activity that inform how information is represented.

In the first two studies, we investigated proactive preparation, exploring how anticipated information is represented in the brain, and whether and how it changes across relevance and probability contexts. The first study focused on the temporal dynamics of preparatory activity. For this purpose, brain activity was recorded with Electroencephalography (EEG). The results showed that several events take place during the preparatory window. First, the perceptual characteristics of the signal are processed. Then, during most of the trial, global differences associated with the mechanisms deployed during attention and expectation were found. Finally, the anticipatory category was represented throughout the trial, especially towards the end. Crucially, although both attention and expectation showed similar decoding accuracies and magnification effects, the results of one did not generalize to those of the other. This suggests that although anticipatory information is represented in the two conditions, the nature of such representations differs. Finally, we sought to explore the level of similarity between preparatory representations and stimulus perception. The two conditions differed in the similarity of patterns between the cue and the target; specifically, the preparation during attention was more similar to the target activity.

This first experiment left an important question open: what are the neural substrates underlying the representations encountered during anticipation? We addressed this question in the second experiment, in which we adapted the behavioral paradigm to functional magnetic resonance imaging (fMRI). Again, we found evidence of anticipatory encoding that was different for attention and expectation. The results revealed increased activations for attention in early visual regions, whereas the posterior cingulate cortex was more active for expectation in this window. In addition, anticipatory information showed large differences in the involvement of different visual and frontoparietal regions during anticipation, suggesting that different networks exhibit distinguishable mechanisms in attention and expectation. In addition, we found data indicative of category representation in the ventral visual cortex, partially replicating the results of the EEG experiment. We found that the attended and probable categories were represented in nearby, though mostly different, regions. Again, cross-classification was not possible between attended and expectation, implying that the representational nature of preparation in the two conditions was different. When comparing the similarity between anticipatory representations and stimulus perception, we found that these similarities occurred in different regions depending on the condition. In addition, a voxel selectivity analysis revealed that only expectation showed some evidence of cue-to-target generalization, and it did so in the ventral visual cortex. Finally, the representation of anticipated stimuli did not correlate with behavior in either condition.

Finally, after studying the nature of top-down preparation, we investigated how the accuracy of prior induced representations changed through attention and expectation. Specifically, we set out to study whether cues sharpen or dampen sensory representations. To do so, we re-analyzed the data from the two experiments to draw joint conclusions. As a preliminary step, we set out to find evidence of separate activation differences for attention and expectation. These analyses showed converging evidence of increased activations for both likely and relevant stimuli. Regarding sharpening and dampening effects, the results mostly supported a sharpening effect for attention, and a marked dampening in expectation. In addition, we examined whether the dampened representations would generalize better to the preparatory interval, as they should be poorer in both cases. Our hypothesis held true, as expectation trials generalized to the signal, whereas relevant targets did not. This again suggests sharpening for attended stimuli and damping for probable stimuli. In the fMRI experiment we studied the involvement of frontoparietal regions in the sharpening and damping effects. Although we were able to successfully decode target categories in some frontoparietal areas, these regions showed no evidence of sharpening or damping for either condition. Finally, we concluded that the results were not consistent with equivalent effects for attention and expectation, suggesting that not only during anticipation but also during target processing relevance and probability appear to act through different mechanisms.

Taken together, our results show that preparation is a complex event, in which the representation of different stimuli is highly dependent on the context in which they occur. This preparation, moreover, affects the fidelity with which perceived stimuli are represented, which is also context-dependent. Finally, this thesis highlights the differences between selective attention and expectation, two processes that, although sharing similarities, involve different mechanisms and manners of representing information.

REFERENCES

Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychological Bulletin*, 142(7), 693.

Aitken, F., & Kok, P. (2022). Hippocampal representations switch from errors to predictions during acquisition of predictive associations. *Nature Communications*, 13(1), 1–13. https://doi.org/10.1038/s41467-022-31040-w

Alilović, J., Timmermans, B., Reteig, L. C., van Gaal, S., & Slagter, H. A. (2019). No Evidence that Predictions and attention Modulate the First Feedforward Sweep of Cortical Information Processing. *Cerebral Cortex*, 1–18. https://doi.org/10.1093/cercor/bhz038

Alink, A., & Blank, H. (2021). Can expectation suppression be explained by reduced attention to predictable stimuli? *Neuroimage*, 231, 117824. https://doi.org/10.1016/j.neuroimage.2021.117824

Alm, K. H., Rolheiser, T., & Olson, I. R. (2016). Inter-individual variation in fronto-temporal connectivity predicts the ability to learn different types of associations. *NeuroImage*, 132, 213–224. https://doi.org/10.1016/j.neuroimage.2016.02.038

Amado, C., Kovács, P., Mayer, R., Ambrus, G. G., Trapp, S., & Kovács, G. (2018). Neuroimaging results suggest the role of prediction in cross-domain priming. *Scientific Reports*, 8(1), 1–10. https://doi.org/10.1038/s41598-018-28696-0

Amunts, K. (2020). Julich-Brain: A 3D probabilistic atlas of the human brain's cytoarchitecture. *Science*. https://www.science.org/doi/10.1126/science.abb4588

Aranda, C., Madrid, E., Tudela, P., & Ruz, M. (2010). Category expectations: A differential modulation of the N170 potential for faces and words. *Neuropsychologia*, 48(14), 4038–4045. https://doi.org/10.1016/j.neuropsychologia.2010.10.002

Auksztulewicz, R., & Friston, K. (2015). Attentional Enhancement of Auditory Mismatch Responses: A DCM/MEG Study. *Cerebral Cortex*, 25(11), 4273–4283. https://doi.org/10.1093/cercor/bhu323

Auksztulewicz, R., & Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex*, 80, 125–140. https://doi.org/10.1016/j.cortex.2015.11.024

Auksztulewicz, R., Friston, K. J., & Nobre, A. C. (2017). Task relevance modulates the behavioural and neural effects of sensory predictions. *PLoS Biology*, 15(12), 1–27. https://doi.org/10.1371/journal.pbio.2003143

Auksztulewicz, R., Schwiedrzik, C. M., Thesen, T., Doyle, W., Devinsky, O., Nobre, A. C., Schroeder, C. E., Friston, K. J., & Melloni, L. (2018). Not all predictions are equal: 'What' and 'When' predictions modulate activity in auditory cortex through different mechanisms. *The Journal of Neuroscience*, 38(40), 0369–18. https://doi.org/10.1523/JNEUROSCI.0369-18.2018

Badre, D., Bhandari, A., Keglovits, H., & Kikumoto, A. (2021). The dimensionality of neural representations for control. *Current Opinion in Behavioral Sciences*, 38, 20–28. https://doi.org/10.1016/j.cobeha.2020.07.002

Bagherzadeh, Y., Baldauf, D., Pantazis, D., & Desimone, R. (2020). Alpha Synchrony and the Neurofeedback Control of Spatial attention. *Neuron*, 105(3), 577-587.e5. https://doi.org/10.1016/j.neuron.2019.11.001

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

Baldauf, D., & Desimone, R. (2014). Neural mechanisms of object-based attention. *Science*, 344(6182), 424–427. https://doi.org/10.1126/science.1247003

Banerjee, S., Snyder, A. C., Molholm, S., & Foxe, J. J. (2011). Oscillatory Alpha-Band Mechanisms and the Deployment of Spatial attention to Anticipated Auditory and Visual Target Locations: Supramodal or Sensory-Specific Control Mechanisms? *Journal of Neuroscience*, 31(27), 9923–9932. https://doi.org/10.1523/JNEUROSCI.4660-10.2011

Barbosa, J., Stein, H., Martinez, R. L., Galan-Gadea, A., Li, S., Dalmau, J., Adam, K. C. S., Valls-Solé, J., Constantinidis, C., & Compte, A. (2020). Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nature Neuroscience*, 23(8), 1016–1024. https://doi.org/10.1038/s41593-020-0644-4

Barron, H. C., Auksztulewicz, R., & Friston, K. (2020). Prediction and memory: A predictive coding account. *Progress in Neurobiology*, 192, 101821. https://doi.org/10.1016/j.pneurobio.2020.101821

Barton, J. J. S., Kuzin, A., Polli, F., & Manoach, D. S. (2006). The use of working memory for task prediction: What benefits accrue from different types of foreknowledge? *Neuroscience*, 139(1), 385–392. https://doi.org/10.1016/j.neuroscience.2005.10.001

Battistoni, E., Stein, T., & Peelen, M. V. (2017). Preparatory attention in visual cortex. *Annals of the New York Academy of Sciences*, 1396(1), 92–107. https://doi.org/10.1111/nyas.13320

Beck, D. M., & Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nature Neuroscience*, 8(8), Article 8. https://doi.org/10.1038/nn1501

Bell, A. H., Summerfield, C., Morin, E. L., Malecek, N. J., & Ungerleider, L. G. (2016). Encoding of stimulus probability in macaque inferior temporal cortex. *Current Biology*. 26(17), 2280–2290. https://doi.org/10.1016/j.cub.2016.07.007

Bichot, N., Rossi, A., & Desimone, R. (2005). Parallel and Serial Neural Mechanisms for Visual Search in Macaque Area V4. *Science.* https://www.science.org/doi/epdf/10.1126/science.1109676?src=getftr&

Blank, H., & Davis, M. H. (2016). Prediction Errors but Not Sharpened Signals Simulate Multivoxel fMRI Patterns during Speech Perception. *PLOS Biology*, 14(11), e1002577. https://doi.org/10.1371/journal.pbio.1002577

Bledowski, C., Kaiser, J., & Rahm, B. (2010). Basic operations in working memory: Contributions from functional imaging studies. *Behavioural Brain Research*, 214(2), 172–179. https://doi.org/10.1016/j.bbr.2010.05.041

Boag, R. J., Stevenson, N., van Dooren, R., Trutti, A. C., Sjoerds, Z., & Forstmann, B. U. (2021). Cognitive Control of Working Memory: A Model-Based Approach. *Brain Sciences*, 11(6), 721. https://doi.org/10.3390/brainsci11060721

Braem, S., & Egner, T. (2018). Getting a Grip on Cognitive Flexibility. *Current Directions in Psychological Science*, 27(6), 470–476. https://doi.org/10.1177/0963721418787475

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436. https://doi.org/10.1163/156856897X00357

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

Brunet, N. M., Bosman, C. A., Vinck, M., Roberts, M., Oostenveld, R., Desimone, R., De Weerd, P., & Fries, P. (2014). Stimulus repetition modulates gamma-band synchronization in primate visual cortex. *Proceedings of the National Academy of Sciences*, 111(9), 3626–3631. https://doi.org/10.1073/pnas.1309714111

Brunia, C. H. M. (1999). Neural aspects of anticipatory behavior. *Acta Psychologica*, 101(2), 213–242. https://doi.org/10.1016/S0001-6918(99)00006-2

Bueno, F. D., & Cravo, A. M. (2021). Post-interval EEG activity is related to task-goals in temporal discrimination. *PLoS ONE*, 16(9 September), 1–19. https://doi.org/10.1371/journal.pone.0257378

Burrows, B. E., Zirnsak, M., Akhlaghpour, H., Wang, M., & Moore, T. (2014). Global selection of saccadic target features by neurons in area v4. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 34(19), 6700–6706. https://doi.org/10.1523/JNEUROSCI.0867-13.2014

Carlson, T. A., Grootswagers, T., & Robinson, A. K. (2019). An introduction to time-resolved decoding analysis for M/EEG.

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal Theta Reflects Uncertainty and Unexpectedness during Exploration and Exploitation. *Cerebral Cortex*, 22(11), 2575–2586. https://doi.org/10.1093/cercor/bhr332

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, 18(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Chawla, D., Rees, G., & Friston, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nature Neuroscience*, 2(7), 671–676. https://doi.org/10.1038/10230

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, 80(6), 2918–2940. https://doi.org/10.1152/jn.1998.80.6.2918

Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, 363(6427), Article 6427. https://doi.org/10.1038/3633450

Christophel, T. B., Cichy, R. M., Hebart, M. N., & Haynes, J. D. (2015). Parietal and early visual cortices encode working memory content across mental transformations. *NeuroImage*, 106, 198–206. https://doi.org/10.1016/j.neuroimage.2014.11.018

Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia*, 49(6), 1407–1409. https://doi.org/10.1016/j.neuropsychologia.2011.01.029

Cichy, R. M., Heinzle, J., & Haynes, J.-D. (2012). Imagery and Perception Share Cortical Representations of Content and Location. *Cerebral Cortex*, 22(2), 372–380. https://doi.org/10.1093/cercor/bhr106

Cichy, R. M., & Oliva, A. (2020). A M/EEG-fMRI Fusion Primer: Resolving Human Brain Responses in Space and Time. *Neuron*, 107(5), 772–781. https://doi.org/10.1016/j.neuron.2020.07.001

Cohen, M. R., & Maunsell, J. H. R. (2011). When attention Wanders: How Uncontrolled Fluctuations in attention Affect Performance. *Journal of Neuroscience*, 31(44), 15802–15806. https://doi.org/10.1523/JNEUROSCI.3063-11.2011

Corrigan, N. M., Richards, T., Webb, S. J., Murias, M., Merkle, K., Kleinhans, N. M., Johnson, L. C., Poliakov, A., Aylward, E., & Dawson, G. (2009). An Investigation of the Relationship Between fMRI and ERP Source Localized Measurements of Brain Activity during Face Processing. *Brain Topography*, 22(2), 83–96. https://doi.org/10.1007/s10548-009-0086-5

Dalski, A., Kovács, G., & Ambrus, G. G. (2022). Evidence for a General Neural Signature of Face Familiarity. *Cerebral Cortex*, 32(12), 2590–2601. https://doi.org/10.1093/cercor/bhab366

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do expectations Shape Perception? *Trends in Cognitive Sciences*, 22(9), 764–779. https://doi.org/10.1016/j.tics.2018.06.002

de Lange, F. P., Rahnev, D. A., Donner, T. H., & Lau, H. (2013). Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *Journal of Neuroscience*, 33(4), 1400–1410. https://doi.org/10.1523/JNEUROSCI.1094-12.2013

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual attention. *Annual Review of Neuroscience*, 18, 193–222.

D'Esposito, M., & Postle, B. R. (2015). The Cognitive Neuroscience of Working Memory. Annual Review of Psychology, 66, 115–142. https://doi.org/10.1146/annurev-psych-010814-015031

Diamond, A. (2013). Executive Functions. Annual Review of Psychology, 64(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dijkstra, N., Bosch, S. E., & Gerven, M. A. J. van. (2017). Vividness of Visual Imagery Depends on the Neural Overlap with Perception in Visual Areas. *Journal of Neuroscience*, 37(5), 1367–1373. https://doi.org/10.1523/JNEUROSCI.3022-16.2016

Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2019). Shared Neural Mechanisms of Visual Perception and Imagery. *Trends in Cognitive Sciences*, 23(5), 423–434. https://doi.org/10.1016/j.tics.2019.02.004

Dijkstra, N., Kok, P., & Fleming, S. M. (2022). Imagery adds stimulus-specific sensory evidence to perceptual detection. *Journal of Vision*, 22(2), 11. https://doi.org/10.1167/jov.22.2.11

Dijkstra, N., Mostert, P., Lange, F. P. de, Bosch, S., & van Gerven, M. A. (2018). Differential temporal dynamics during visual imagery and perception. *eLife*, 7, e33904. https://doi.org/10.7554/eLife.33904

Dodds, C. M., Morein-Zamir, S., & Robbins, T. W. (2011). Dissociating Inhibition, attention, and Response Control in the Frontoparietal Network Using Functional Magnetic Resonance Imaging. *Cerebral Cortex*, 21(5), 1155–1165. https://doi.org/10.1093/cercor/bhq187

Doesburg, S. M., Roggeveen, A. B., Kitajo, K., & Ward, L. M. (2008). Large-Scale Gamma-Band Phase Synchronization and Selective attention. *Cerebral Cortex*, 18(2), 386–396. https://doi.org/10.1093/cercor/bhm073

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

Dosenbach, N. U. F., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A. T., Fox, M. D., Snyder, A. Z., Vincent, J. L., Raichle, M. E., Schlaggar, B. L., & Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences*, 104(26), 11073–11078. https://doi.org/10.1073/pnas.0704320104

Dumas, J. A. (2015). What Is Normal Cognitive Aging? Evidence from Task-Based Functional Neuroimaging. *Current Behavioral Neuroscience Reports*, 2(4), 256–261. https://doi.org/10.1007/s40473-015-0058-x

Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *NeuroImage*, 39(2), 647–660. https://doi.org/10.1016/j.neuroimage.2007.09.034

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, 14(4), 172–179. https://doi.org/10.1016/j.tics.2010.01.004

Eger, E., Henson, R., Driver, J., & Dolan, R. (2007). Mechanisms of Top-Down Facilitation in Perception of Visual Objects Studied by fMRI. *Cerebral Cortex*, 17(9), 2123–2133. https://doi.org/10.1093/cercor/bhl119

Egner, T., Monti, J. M., & Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *Journal of Neuroscience*, 30(49), 16601–16608. https://doi.org/10.1523/JNEUROSCI.2770-10.2010

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, 113(28), 7000–7005. https://doi.org/10.1073/pnas.1602413113

Ekman, M., Kok, P., & De Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications*, 8(May), 1–9. https://doi.org/10.1038/ncomms15276

Esterman, M., Tamber-Rosenau, B. J., Chiu, Y.-C., & Yantis, S. (2010). Avoiding non-independence in fMRI data analysis: Leave one subject out. *NeuroImage*, 50(2), 572–576. https://doi.org/10.1016/j.neuroimage.2009.10.092

Esterman, M., & Yantis, S. (2010). Perceptual expectation evokes category-selective cortical activity. *Cerebral Cortex*, 20(5), 1245–1253. https://doi.org/10.1093/cercor/bhp188

Feldman, H., & Friston, K. (2010). Attention, Uncertainty, and Free-Energy. *Frontiers in Human Neuroscience*, 4. https://www.frontiersin.org/articles/10.3389/fnhum.2010.00215

Feng, W., Störmer, V. S., Martinez, A., McDonald, J. J., & Hillyard, S. A. (2017). Involuntary orienting of attention to a sound desynchronizes the occipital alpha rhythm and improves visual perception. *NeuroImage*, 150, 318–328. https://doi.org/10.1016/j.neuroimage.2017.02.033

Fernández, A., Okun, S., & Carrasco, M. (2022). Differential Effects of Endogenous and Exogenous attention on Sensory Tuning. *Journal of Neuroscience*, 42(7), 1316–1327. https://doi.org/10.1523/JNEUROSCI.0892-21.2021

Feuerriegel, D., Churches, O., Coussens, S., & Keage, H. A. D. (2018). Evidence for spatiotemporally distinct effects of image repetition and perceptual expectations as measured by event-related potentials. *NeuroImage*, 169, 94–105. https://doi.org/10.1016/j.neuroimage.2017.12.029

Feuerriegel, D., Vogels, R., & Kovacs, G. Z. (2021). Evaluating the evidence for expectation suppression in the visual system | Elsevier Enhanced Reader. *Neuroscience and Biovahavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2021.04.002

Feuerriegel, D., Yook, J., Quek, G. L., Hogendoorn, H., & Bode, S. (2021). Visual mismatch responses index surprise signalling but not expectation suppression. Cortex, 134, 16–29. https://doi.org/10.1016/j.cortex.2020.10.006

Ficco, L., Mancuso, L., Manuello, J., Teneggi, A., Liloia, D., Duca, S., Costa, T., Kovacs, G. Z., & Cauda, F. (2021). Disentangling predictive processing in the brain: A meta-analytic study in favour of a predictive network. *Scientific Reports*, 11(1), Article 1. https://doi.org/10.1038/s41598-021-95603-5

Fishman, R. S. (1997). Gordon Holmes, the cortical retina, and the wounds of war. The seventh Charles B. Snyder Lecture. Documenta Ophthalmologica. *Advances in Ophthalmology*, 93(1–2), 9–28. https://doi.org/10.1007/BF02569044

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10), 474–480.

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of Oscillatory Neuronal Synchronization by Selective Visual attention. *Science*, 291(5508), 1560–1563. https://doi.org/10.1126/science.1055465

Friston, K. (2005). A theory of cortical responses. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 360(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. Clinical Neurophysiology, 120(3), 453–463. https://doi.org/10.1016/j.clinph.2008.11.029

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, 16(2), 129–135. https://doi.org/10.1016/j.tics.2011.11.014

Giesbrecht, B., Weissman, D. H., Woldorff, M. G., & Mangun, G. R. (2006). Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Research*, 1080(1), 63–72. https://doi.org/10.1016/j.brainres.2005.09.068

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., & Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615), Article 7615. https://doi.org/10.1038/nature18933

Goddard, E., Carlson, T. A., & Woolgar, A. (2022). Spatial and Feature-selective attention Have Distinct, Interacting Effects on Population-level Tuning. *Journal of Cognitive Neuroscience*, 34(2), 290–312. https://doi.org/10.1162/jocn_a_01796

González-García, C., Arco, J. E., Palenciano, A. F., Ramírez, J., & Ruz, M. (2017). Encoding, preparation and implementation of novel complex verbal instructions. *NeuroImage*, 148, 264–273. https://doi.org/10.1016/j.neuroimage.2017.01.037

González-García, C., Flounders, M. W., Chang, R., Baria, A. T., & He, B. J. (2018). Content-specific activity in frontoparietal and default-mode networks during prior-guided visual perception. *eLife*, 7, 1–25. https://doi.org/10.7554/eLife.36068

González-García, C., Formica, S., Wisniewski, D., & Brass, M. (2021). Frontoparietal action-oriented codes support novel instruction implementation. *NeuroImage*, 226(October 2020). https://doi.org/10.1016/j.neuroimage.2020.117608

González-García, C., & He, B. J. (2021). A Gradient of Sharpening Effects by Perceptual Prior across the Human Cortical Hierarchy. *Journal of Neuroscience,* 41(1), 167–178.

González-García, C., Mas-Herrero, E., de Diego-Balaguer, R., & Ruz, M. (2016). Task-specific preparatory neural activations in low-interference contexts. Brain Structure and Function, 221(8), 3997–4006. https://doi.org/10.1007/s00429-015-1141-5

Gordon, N., Tsuchiya, N., Koenig-Robert, R., & Hohwy, J. (2019). Expectation and attention increase the integration of top-down and bottom-up signals in perception through different pathways. *PLOS Biology*, 17(4), e3000233. https://doi.org/10.1371/journal.pbio.3000233

Greenberg, A. S., Esterman, M., Wilson, D., Serences, J. T., & Yantis, S. (2010). Control of Spatial and Feature-Based attention in Frontoparietal Cortex. *Journal of Neuroscience*, 30(43), 14330–14339. https://doi.org/10.1523/JNEUROSCI.4248-09.2010

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, 41(10), 1409–1422. https://doi.org/10.1016/S0042-6989(01)00073-6

Grootswagers, T., Wardel, S. G., & Carlson, T. A. (2017). Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data. *Journal of Cognitive Neuroscience*, 29(4), 677–697. https://doi.org/10.1162/jocn_a_01068

Gruber, T., Müller, M. M., Keil, A., & Elbert, T. (1999). Selective visual-spatial attention alters induced gamma band responses in the human EEG. Clinical Neurophysiology, 110(12), 2074–2085. https://doi.org/10.1016/S1388-2457(99)00176-5

Hall, M. G., Mattingley, J. B., & Dux, P. E. (2018). Electrophysiological correlates of incidentally learned expectations in human vision. *Journal of Neurophysiology*, 119(4), 1461–1470. https://doi.org/10.1152/jn.00733.2017

Han, B., Mostert, P., & de Lange, F. P. (2019). Predictable tones elicit stimulus-specific suppression of evoked activity in auditory cortex. *NeuroImage*, 200, 242–249. https://doi.org/10.1016/j.neuroimage.2019.06.033

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* (New York, N.Y.), 293(5539), 2425–2430. https://doi.org/10.1126/science.1063736

Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding Neural Representational Spaces Using Multivariate Pattern Analysis. *Annual Review of Neuroscience*, 37(1), 435–456. https://doi.org/10.1146/annurev-neuro-062012-170325

He, X., Fan, S., Zhou, K., & Chen, L. (2004). Cue Validity and Object-Based attention. *Journal of Cognitive Neuroscience*, 16(6), 1085–1097. https://doi.org/10.1162/0898929041502689

Hebart, M. N., & Baker, C. I. (2018). Deconstructing multivariate decoding for the study of brain function. *NeuroImage*, 180(August 2017), 4–18. https://doi.org/10.1016/j.neuroimage.2017.08.005

Hebart, M. N., Bankson, B. B., Harel, A., Baker, C. I., & Cichy, R. M. (2018). The representational dynamics of task and object processing in humans. *eLife*, 7, e32816. https://doi.org/10.7554/eLife.32816

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., & Rugg, M. D. (2003). Electrophysiological and Haemodynamic Correlates of Face Perception, Recognition and Priming. *Cerebral Cortex*, 13(7), 793–805. https://doi.org/10.1093/cercor/13.7.793

Hindy, N. C., Ng, F. Y., & Turk-Browne, N. B. (2016). Linking pattern completion in the hippocampus to predictive coding in visual cortex. *Nature Neuroscience*, 19(5), 665–667. https://doi.org/10.1038/nn.4284

Ho, T. C., Brown, S., Abuyo, N. A., Ku, E. H. J., & Serences, J. T. (2012). Perceptual consequences of feature-based attentional enhancement and suppression. *Journal of Vision*, 12(8), 1–17. https://doi.org/10.1167/12.8.15

Hommel, B., Chapman, C. S., Cisek, P., Neyedli, H. F., Song, J.-H., & Welsh, T. N. (2019). No one knows what attention is. *Attention, Perception, & Psychophysics*, 81(7), 2288–2303. https://doi.org/10.3758/s13414-019-01846-w

Hong, X., Wang, Y., Sun, J., Li, C., & Tong, S. (2017). Segregating Top-Down Selective attention from Response Inhibition in a Spatial Cueing Go NoGo Task: An ERP and Source Localization Study. *Scientific Reports*, 7(1), 1–14. https://doi.org/10.1038/s41598-017-08807-z

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3), 574–591.

Jafarpour, A., Horner, A. J., Fuentemilla, L., Penny, W. D., & Duzel, E. (2013). Decoding oscillatory representations and mechanisms in memory. *Neuropsychologia*, 51(4), 772–780. https://doi.org/10.1016/j.neuropsychologia.2012.04.002

James, W. (2007). The Principles of Psychology. Cosimo, Inc.

Jiang, J., Summerfield, C., & Egner, T. (2013). Attention Sharpens the Distinction between Expected and Unexpected Percepts in the Visual Brain. *Journal of Neuroscience*, 33(47), 18438–18447. https://doi.org/10.1523/JNEUROSCI.3308-13.2013

Jin, W., Nobre, A. C., & van Ede, F. (2020). Temporal expectations prepare visual working memory for behavior. *Journal of Cognitive Neuroscience*, 32(12), 2320–2332. https://doi.org/10.1162/jocn_a_01626

Jokisch, D., & Jensen, O. (2007). Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *The Journal of Neuroscience*: The Official Journal of the Society for Neuroscience, 27(12), 3244–3251. https://doi.org/10.1523/JNEUROSCI.5399-06.2007

Junghöfer, M., Elbert, T., Tucker, D. M., & Braun, C. (1999). The polar average reference effect: A bias in estimating the head surface integral in EEG recording. *Clinical Neurophysiology*, 110(6), 1149–1155. https://doi.org/10.1016/S1388-2457(99)00044-9

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *Journal of Neuroscience*, 17(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Kaplan, J. T., Man, K., & Greening, S. G. (2015). Multivariate cross-classification: Applying machine learning techniques to characterize abstraction in neural representations. *Frontiers in Human Neuroscience*, 9(March), 1–12. https://doi.org/10.3389/fnhum.2015.00151

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, 22(4), 751–761. https://doi.org/10.1016/S0896-6273(00)80734-5

Kastner, S., Weerd, P. D., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of Directed attention in the Human Extrastriate Cortex as Revealed by Functional MRI. *Science*, 282, 108–111.

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., Luu, P., Miller, G. A., & Yee, C. M. (2014). Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology*, 51(1), 1–21. https://doi.org/10.1111/psyp.12147

Kerrén, C., Linde-Domingo, J., Hanslmayr, S., & Wimber, M. (2018). An Optimal Oscillatory Phase for Pattern Reactivation during Memory Retrieval. *Current Biology*, 28(21), 3383-3392.e6. https://doi.org/10.1016/j.cub.2018.08.065

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, 8(3), 159–166. https://doi.org/10.1007/s10339-007-0170-2

Kim, H., Schlichting, M. L., Preston, A. R., & Lewis-Peacock, J. A. (2020). Predictability changes what we remember in familiar temporal contexts. *Journal of Cognitive Neuroscience*, 32(1), 124–140. https://doi.org/10.1162/jocn_a_01473

Kim, N. Y., & Kastner, S. (2019). A biased competition theory for the developmental cognitive neuroscience of visuo-spatial attention. *Current Opinion in Psychology*, 29, 219–228. https://doi.org/10.1016/j.copsyc.2019.03.017

King, J. R., Faugeras, F., Gramfort, A., Schurger, A., El Karoui, I., Sitt, J. D., Rohaut, B., Wacongne, C., Labyt, E., Bekinschtein, T., Cohen, L., Naccache, L., & Dehaene, S. (2013). Single-trial decoding of auditory novelty responses facilitates the detection of residual consciousness. *NeuroImage*, 83, 726–738. https://doi.org/10.1016/j.neuroimage.2013.07.013

King, J.-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences*, 18(4), 203–210. https://doi.org/10.1016/j.tics.2014.01.002

Koenig-Robert, R., & Pearson, J. (2021). Why do imagery and perception look and feel so different? Philosophical Transactions of the Royal Society of London. Series B, *Biological Sciences*, 376(1817), 20190703. https://doi.org/10.1098/rstb.2019.0703

Kok, P., & de Lange, F. P. (2015). Predictive Coding in Sensory Cortex. In B. U. Forstmann & E.-J. Wagenmakers (Eds.), An Introduction to Model-Based Cognitive Neuroscience (pp. 221–244). Springer. https://doi.org/10.1007/978-1-4939-2236-9_11

Kok, P., Failing, M. F., & de Lange, F. P. (2014). Prior expectations Evoke Stimulus Templates in the Primary Visual Cortex. *Journal of Cognitive Neuroscience*, 26(7), 1546–1554. https://doi.org/10.1162/jocn_a_00562

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less Is More: expectation Sharpens Representations in the Primary Visual Cortex. *Neuron*, 75(2), 265–270. https://doi.org/10.1016/j.neuron.2012.04.034

Kok, P., Lieshout, L. L. F. van, & Lange, F. P. de. (2016). Local expectation violations result in global activity gain in primary visual cortex. *Scientific Reports*, 6. https://doi.org/10.1038/srep37706

Kok, P., Mostert, P., & de Lange, F. P. (2017). Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences*, 201705652. https://doi.org/10.1073/pnas.1705652114

Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & De Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, 22(9), 2197–2206. https://doi.org/10.1093/cercor/bhr310

Kok, P., & Turk-Browne, N. B. (2018). Associative Prediction of Visual Shape in the Hippocampus. *The Journal of Neuroscience*, 38(31), 6888–6899. https://doi.org/10.1523/JNEUROSCI.0163-18.2018

Koshino, H., Minamoto, T., Yaoi, K., Osaka, M., & Osaka, N. (2015). Coactivation of the Default Mode Network regions and Working Memory Network regions during task preparation. *Scientific Reports*, 4(1), 5954. https://doi.org/10.1038/srep05954

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational Similarity Analysis – Connecting the Branches of Systems Neuroscience. *Frontiers in Systems Neuroscience*, 2, 4. https://doi.org/10.3389/neuro.06.004.2008

Kumar, S., Kaposvari, P., & Vogels, R. (2017). Encoding of Predictable and Unpredictable Stimuli by Inferior Temporal Cortical Neurons. *Journal of Cognitive Neuroscience*, 29(8), 1445–1454. https://doi.org/10.1162/jocn_a_01135

Kveraga, K., Ghuman, A. S., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition*, 65(2), 145–168. https://doi.org/10.1016/j.bandc.2007.06.007

Langers, D. R. M., & van Dijk, P. (2012). Mapping the Tonotopic Organization in Human Auditory Cortex with Minimally Salient Acoustic Stimulation. *Cerebral Cortex* (New York, NY), 22(9), 2024–2038. https://doi.org/10.1093/cercor/bhr282

Lawrence, S. J. D., Formisano, E., Muckli, L., & de Lange, F. P. (2019). Laminar fMRI: Applications for cognitive neuroscience. *NeuroImage*, 197, 785–791. https://doi.org/10.1016/j.neuroimage.2017.07.004

Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2013). Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nature Neuroscience*, 16(8), Article 8. https://doi.org/10.1038/nn.3452

Litwin, P., & Miłkowski, M. (2020). Unification by Fiat: Arrested Development of Predictive Processing. *Cognitive Science*, 44(7), e12867. https://doi.org/10.1111/cogs.12867

Liu, T., Stevens, S. T., & Carrasco, M. (2007). Comparing the time course and efficacy of spatial and feature-based attention. Vision Research, 47(1), 108–113. https://doi.org/10.1016/j.visres.2006.09.017

López-García, D., Peñalver, J. M. G., Górriz, J. M., & Ruz, M. (2022). MVPAlab: A machine learning decoding toolbox for multidimensional electroencephalography data. *Computer Methods and Programs in Biomedicine*, 214, 106549. https://doi.org/10.1016/j.cmpb.2021.106549

López-García, D., Sobrado, A., Peñalver, J. M. G., Górriz, J. M., & Ruz, M. (2020). Multivariate Pattern Analysis Techniques for Electroencephalography Data to Study Flanker Interference Effects. International Journal of

Neural Systems, 30(7). https://doi.org/10.1142/S0129065720500240

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J. N., Morey, R. D., & Wagenmakers, E. J. (2019). JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software*, 88(1). https://doi.org/10.18637/jss.v088.i02

Lowe, M., Teng, S., Mohsenzadeh, Y., Charest, I., Pantazis, D., & Oliva, A. (2018). Temporal dynamics underlying sound discrimination in the human brain. 2018 Conference on Cognitive Computational Neuroscience. 2018 Conference on Cognitive Computational Neuroscience, Philadelphia, Pennsylvania, USA. https://doi.org/10.32470/CCN.2018.1090-0

Luck, S. J., & Kappenman, E. S. (2011). The Oxford Handbook of Event-Related Potential Components. Oxford University Press.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, 47(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Mackie, M.-A., Van Dam, N. T., & Fan, J. (2013). Cognitive Control and attentional Functions. *Brain and Cognition*, 82(3), 301–312. https://doi.org/10.1016/j.bandc.2013.05.004

Magri, C., Schridde, U., Murayama, Y., Panzeri, S., & Logothetis, N. K. (2012). The Amplitude and Timing of the BOLD Signal Reflects the Relationship between Local Field Potential Power at Different Frequencies. *The Journal of Neuroscience*, 32(4), 1395–1407. https://doi.org/10.1523/JNEUROSCI.3985-11.2012

Manahova, M. E., Mostert, P., Kok, P., Schoffelen, J.-M., & de Lange, F. P. (2018). Stimulus Familiarity and expectation Jointly Modulate Neural Activity in the Visual Ventral Stream. *Journal of Cognitive Neuroscience*, 30(9), 1366–1377. https://doi.org/10.1162/jocn_a_01281

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-Based attention Increases the Selectivity of Population Responses in Primate Visual Cortex. *Current Biology*, 14(9), 744–751. https://doi.org/10.1016/j.cub.2004.04.028

Mechelli, A., Humphreys, G. W., Mayall, K., Olson, A., & Price, C. J. (2000). Differential effects of word length and visual contrast in the fusiform and lingual gyri during. Proceedings of the Royal Society of London. Series B: Biological Sciences, 267(1455), 1909–1913. https://doi.org/10.1098/rspb.2000.1229

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable

Analysis. Cognitive Psychology, 41(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Molinaro, N., Barraza, P., & Carreiras, M. (2013). Long-range neural synchronization supports fast and efficient reading: EEG correlates of processing expected words in sentences. *NeuroImage*, 72, 120–132. https://doi.org/10.1016/j.neuroimage.2013.01.031

Mongillo, G., Barak, O., & Tsodyks, M. (2008). SynaptiC Theory of Working Memory. *Science*, 319(5869), 1543–1546. https://doi.org/10.1126/science.1150769

Mostert, P., Albers, A. M., Brinkman, L., Todorova, L., Kok, P., & de Lange, F. P. (2018). Eye Movement-Related Confounds in Neural Decoding of Visual Working Memory Representations. *eNeuro*, 5(4), ENEURO.0401-17.2018. https://doi.org/10.1523/ENEURO.0401-17.2018

Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., Goebel, R., & Yacoub, E. (2015). Contextual Feedback to Superficial Layers of V1. *Current Biology*: CB, 25(20), 2690–2695. https://doi.org/10.1016/j.cub.2015.08.057

Mueller, V., Brehmer, Y., von Oertzen, T., Li, S.-C., & Lindenberger, U. (2008). Electrophysiological correlates of selective attention: A lifespan comparison. *BMC Neuroscience*, 9(1), 18. https://doi.org/10.1186/1471-2202-9-18

Natale, E., Marzi, C. A., Girelli, M., Pavone, E. F., & Pollmann, S. (2006). ERP and fMRI correlates of endogenous and exogenous focusing of visual-spatial attention. European *Journal of Neuroscience*, 23(9), 2511–2521. https://doi.org/10.1111/j.1460-9568.2006.04756.x

Nichols, T. E., Das, S., Eickhoff, S. B., Evans, A. C., Glatard, T., Hanke, M., Kriegeskorte, N., Milham, M. P., Poldrack, R. A., Poline, J.-B., Proal, E., Thirion, B., Essen, D. C. V., White, T., & Yeo, B. T. T. (2016). Best Practices in Data Analysis and Sharing in Neuroimaging using MRI (p. 054262). bioRxiv. https://doi.org/10.1101/054262

Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A Toolbox for Representational Similarity Analysis. *PLoS Computational Biology*, 10(4). https://doi.org/10.1371/journal.pcbi.1003553

Nir-Cohen, G., Kessler, Y., & Egner, T. (2020). Neural Substrates of Working Memory Updating. *Journal of Cognitive Neuroscience*, 32(12), 2285–2302. https://doi.org/10.1162/jocn_a_01625

Nobre, A. C., & Serences, J. T. (2018). Building on a Solid Baseline: Anticipatory Biases in attention. *Trends in Neurosciences*, 41(3), 120–122. https://doi.org/10.1016/j.tins.2018.01.005

Norman, D. A., & Shallice, T. (1986). Attention to Action. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), Consciousness and Self-Regulation: Advances in Research and Theory Volume 4 (pp. 1–18). Springer US. https://doi.org/10.1007/978-1-4757-0629-1_1

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430. https://doi.org/10.1016/j.tics.2006.07.005

Novitskiy, N., Ramautar, J. R., Vanderperren, K., De Vos, M., Mennes, M., Mijovic, B., Vanrumste, B., Stiers, P., Van den Bergh, B., Lagae, L., Sunaert, S., Van Huffel, S., & Wagemans, J. (2011). The BOLD correlates of the visual P1 and N1 in single-trial analysis of simultaneous EEG-fMRI recordings during a spatial detection task. *NeuroImage*, 54(2), 824–835. https://doi.org/10.1016/j.neuroimage.2010.09.041

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34(1), 171–175. https://doi.org/10.1016/0006-8993(71)90358-1

Oostenveld, R., Fries, P., Mars, E., & Schoffelen, J. M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. https://www.hindawi.com/journals/cin/2011/156869/

Ortiz-Tudela, J., Bergmann, J., Bennett, M., Ehrlich, I., Muckli, L., & Shing, Y. L. (2023). Concurrent contextual and time-distant mnemonic information co-exist as feedback in the human visual cortex. *Neuroimage*, 265, 119778. https://doi.org/10.1016/j.neuroimage.2022.119778

Ortiz-Tudela, J., & González-García, C. (2023). Exploring the conceptual structure of Spanish Experimental Psychology. https://doi.org/10.20350/digitalCSIC/15275

Pace, T., Koenig-Robert, R., & Pearson, J. (2023). Different Mechanisms for Supporting Mental Imagery and Perceptual Representations: Modulation Versus Excitation. Psychological Science, 09567976231198435. https://doi.org/10.1177/09567976231198435

Palenciano, A. F., González-García, C., Arco, J. E., Pessoa, L., & Ruz, M. (2019). Representational Organization of Novel Task Sets during Proactive Encoding. *Journal of Neuroscience*, 39(42), 8386–8397. https://doi.org/10.1523/JNEUROSCI.0725-19.2019

Palenciano, A. F., González-García, C., Arco, J. E., & Ruz, M. (2019). Transient and Sustained Control Mechanisms Supporting Novel Instructed Behavior. *Cerebral Cortex*, 29(9), 3948–3960. https://doi.org/10.1093/cercor/bhy273

Palenciano, A. F., Senoussi, M., Formica, S., & González-García, C. (2023). Canonical template tracking: Measuring the activation state of specific neural representations. Frontiers in Neuroimaging, 1. https://www.frontiersin.org/articles/10.3389/fnimg.2022.974927

Panichello, M. F., & Buschman, T. J. (2021). Shared mechanisms underlie the control of working memory and attention. Nature, 592(7855), Article 7855. https://doi.org/10.1038/s41586-021-03390-w

Pazo-Alvarez, P., Cadaveira, F., & Amenedo, E. (2003). MMN in the visual modality: A review. Biological Psychology, 63(3), 199–236. https://doi.org/10.1016/S0301-0511(03)00049-8

Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences* of the United States of America, 108(29), 12125–12130. https://doi.org/10.1073/pnas.1101042108

Pelzer, L., Haffmann, J., Naefgen, C., Gaschler, R., & Haider, H. (2022). Task-separation in dual-tasking: How action effects support the separation of the task streams. *Acta Psychologica*, 222, 103464. https://doi.org/10.1016/j.actpsy.2021.103464

Peñalver, J. M. G., López-García, D., González-García, C., Aguado-López, B., Górriz, J. M., & Ruz, M. (2023). Top-down specific preparatory activations for selective attention and perceptual expectations. *NeuroImage*, 271, 119960. https://doi.org/10.1016/j.neuroimage.2023.119960

Pernet, C., Garrido, M., Gramfort, A., Maurits, N., Michel, C., Pang, E., Salmelin, R., Schoffelen, J. M., Valdes-Sosa, P., & Puce, A. (2018). Best Practices in Data Analysis and Sharing in Neuroimaging using MEEG. 1–64. https://doi.org/10.31219/osf.io/a8dhx

Perrone-Bertolotti, M., El Bouzaïdi Tiali, S., Vidal, J. R., Petton, M., Croize, A. C., Deman, P., Rheims, S., Minotti, L., Bhattacharjee, M., Baciu, M., Kahane, P., & Lachaux, J. P. (2020). A real-time marker of object-based attention in the human brain. A possible component of a "gate-keeping mechanism" performing late attentional selection in the Ventro-Lateral Prefrontal Cortex. *NeuroImage*, 210, 116574. https://doi.org/10.1016/j.neuroimage.2020.116574

Petersen, S. E., & Posner, M. I. (2012). The attention System of the Human Brain: 20 Years After. *Annual Review of Neuroscience*, 35(1), 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525

Petro, L. S., & Muckli, L. (2017). The laminar integration of sensory inputs with feedback signals in human cortex. *Brain and Cognition*, 112, 54–57. https://doi.org/10.1016/j.bandc.2016.06.007

Poldrack, R. (2016, April 18). russpoldrack.org: How folksy is psychology? The linguistic history of cognitive ontologies. Russpoldrack.Org.

https://russpoldrack.blogspot.com/2016/04/how-folksy-is-psychology-linguistic.html

Poldrack, R. A. (2021). The physics of representation. *Synthese*, 199(1), 1307–1325. https://doi.org/10.1007/s11229-020-02793-y

Poldrack, R. A., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., Stott Parker, D., Sabb, F. W., & Bilder, R. M. (2011). The Cognitive Atlas: Toward a Knowledge Foundation for Cognitive Neuroscience. *Frontiers in Neuroimaging*, 5(17), 1–11.

Pooresmaeili, A., Poort, J., & Roelfsema, P. R. (2014). Simultaneous selection by object-based attention in visual and frontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 111(17), 6467–6472. https://doi.org/10.1073/pnas.1316181111

Posner, M. I., & Snyder, C. R. R. (2004). Attention and Cognitive Control. In D. A. Balota & E. J. Marsh (Eds.), *Cognitive psychology: Key readings* (205–223). Psychology Press.

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, 32(1), 3–25. https://doi.org/10.1080/00335558008248231

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

Postle, B. R. (2006). Working Memory as an Emergent Property of the Mind and Brain. *Neuroscience*, 139(1), 23–38. https://doi.org/10.1016/j.neuroscience.2005.06.005

Potts, G. F. (2004). An ERP index of task relevance evaluation of visual stimuli. *Brain and Cognition*, 56(1), 5–13. https://doi.org/10.1016/j.bandc.2004.03.006

Press, C., Kok, P., & Yon, D. (2020). The Perceptual Prediction Paradox. *Trends in Cognitive Sciences*, 24(1), 13–24. https://doi.org/10.1016/j.tics.2019.11.003

Quax, S. C., Dijkstra, N., van Staveren, M. J., Bosch, S. E., & van Gerven, M. A. J. (2019). Eye movements explain decodability during perception and cued attention in MEG. *NeuroImage*, 195, 444–453. https://doi.org/10.1016/j.neuroimage.2019.03.069

Rahnev, D., Maniscalco, B., Graves, T., Huang, E., de Lange, F. P., & Lau, H. (2011). Attention induces conservative subjective biases in visual perception. *Nature Neuroscience*, 14(12), Article 12. https://doi.org/10.1038/nn.2948

Ranganath, C., Cohen, M. X., Dam, C., & D’Esposito, M. (2004). Inferior Temporal, Prefrontal, and Hippocampal Contributions to Visual Working Memory Maintenance and Associative Memory Retrieval.

Journal of Neuroscience, 24(16), 3917–3925. https://doi.org/10.1523/JNEUROSCI.5053-03.2004

Rauss, K., & Pourtois, G. (2013). What is Bottom-Up and What is Top-Down in Predictive Coding? *Frontiers in Psychology*, 4. https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00276

Richter, D., & de Lange, F. P. (2019). Statistical learning attenuates visual activity only for attended stimuli. *eLife*, 8. Scopus. https://doi.org/10.7554/eLife.47869

Richter, D., Ekman, M., & de Lange, F. P. (2018). Suppressed Sensory Response to Predictable Object Stimuli throughout the Ventral Visual Stream. *The Journal of Neuroscience*, 38(34), 7452–7461. https://doi.org/10.1523/JNEUROSCI.3421-17.2018

Rohenkohl, G., Cravo, A. M., Wyart, V., & Nobre, A. C. (2012). Temporal expectation improves the quality of sensory information. *Journal of Neuroscience*, 32(24), 8424–8428. https://doi.org/10.1523/JNEUROSCI.0804-12.2012

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, 354(6316), 1136–1139. https://doi.org/10.1126/science.aah7011

Roth, J. K., Serences, J. T., & Courtney, S. M. (2006). Neural System for Controlling the Contents of Object Working Memory in Humans. *Cerebral Cortex*, 16(11), 1595–1603. https://doi.org/10.1093/cercor/bhj096

Rousselet, G. A., Macé, M. J.-M., & Fabre-Thorpe, M. (2004). Animal and human faces in natural scenes: How specific to human faces is the N170 ERP component? *Journal of Vision*, 4(1), 2. https://doi.org/10.1167/4.1.2

Rungratsameetaweemana, N., Itthipuripat, S., Salazar, A., & Serences, J. T. (2018). Expectations Do Not Alter Early Sensory Processing during Perceptual Decision-Making. *Journal of Neuroscience*, 38(24), 5632–5648. https://doi.org/10.1523/JNEUROSCI.3638-17.2018

Rungratsameetaweemana, N., & Serences, J. T. (2019). Dissociating the impact of attention and expectation on early sensory processing. *Current Opinion in Psychology*, 29, 181–186. https://doi.org/10.1016/j.copsyc.2019.03.014

Rungratsameetaweemana, N., Squire, L. R., & Serences, J. T. (2019). Preserved capacity for learning statistical regularities and directing selective attention after hippocampal lesions. *Proceedings of the National Academy of Sciences*, 116(39), 19705–19710. https://doi.org/10.1073/pnas.1904502116

Ruz, M., & Nobre, A. C. (2008). Attention modulates initial stages of visual word processing. *Journal of Cognitive Neuroscience*, 20(9), 1727–1736. https://doi.org/10.1162/jocn.2008.20119

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience*, 5(7), 631–632. https://doi.org/10.1038/nn876

Scheeringa, R., Fries, P., Petersson, K.-M., Oostenveld, R., Grothe, I., Norris, D. G., Hagoort, P., & Bastiaansen, M. C. M. (2011). Neuronal Dynamics Underlying High- and Low-Frequency EEG Oscillations Contribute Independently to the Human BOLD Signal. *Neuron*, 69(3), 572–583. https://doi.org/10.1016/j.neuron.2010.11.044

Schröger, E. (1996). A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *Journal of Cognitive Neuroscience*, 8(6), 527–539. https://doi.org/10.1162/jocn.1996.8.6.527

Schröger, E., Kotz, S. A., & SanMiguel, I. (2015). Bridging prediction and attention in current research on perception and action. *Brain Research*, 1626, 1–13. https://doi.org/10.1016/j.brainres.2015.08.037

Scolari, M., Byers, A., & Serences, J. T. (2012). Optimal Deployment of attentional Gain during Fine Discriminations. *The Journal of Neuroscience*, 32(22), 7723–7733. https://doi.org/10.1523/JNEUROSCI.5558-11.2012

Self, M. W., van Kerkoerle, T., Goebel, R., & Roelfsema, P. R. (2019). Benchmarking laminar fMRI: Neuronal spiking and synaptic activity during top-down and bottom-up processing in the different layers of cortex. *NeuroImage*. https://pubmed.ncbi.nlm.nih.gov/28648888/

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. Psychological Science, 20(2), 207–214. https://doi.org/10.1111/j.1467-9280.2009.02276.x

Serences, J. T., Schwarzbach, J., Courtney, S. M., Golay, X., & Yantis, S. (2004). Control of Object-based attention in Human Cortex. *Cerebral Cortex*, 14(12), 1346–1357. https://doi.org/10.1093/cercor/bhh095

Simon, A., Schachtner, J., & Gallen, C. L. (2019). Disentangling expectation from selective attention during perceptual decision-making. SSRN Electronic Journal, 5(564), 1–19. https://doi.org/10.4324/9781315853178

Simon, A., Schatner, J., & Gallen, C. (2018). Disentangling expectation from selective attention during perceptual decision making. *Journal of Neurophysiology*. https://journals.physiology.org/doi/full/10.1152/jn.00639.2018

Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. Applied Soft Computing, 97, 105524. https://doi.org/10.1016/j.asoc.2019.105524

Smith, F. W., & Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. *Proceedings of the National Academy of Sciences* of the United States of America, 107(46), 20099–20103. https://doi.org/10.1073/pnas.1000233107

Snyder, A. C., & Foxe, J. J. (2010). Anticipatory attentional Suppression of Visual Features Indexed by Oscillatory Alpha-Band Power Increases:A High-Density Electrical Mapping Study. *The Journal of Neuroscience*, 30(11), 4024–4032.

Soon, C. S., Namburi, P., & Chee, M. W. L. (2013). Preparatory patterns of neural activity predict visual category search speed. *NeuroImage*, 66, 215–222. https://doi.org/10.1016/j.neuroimage.2012.10.036

Soto, D., & Blanco, M. J. (2004). Spatial attention and object-based attention: A comparison within a single task. *Vision Research*, 44(1), 69–81. https://doi.org/10.1016/j.visres.2003.08.013

Sperling, R., Chua, E., Cocchiarella, A., Rand-Giovannetti, E., Poldrack, R., Schacter, D. L., & Albert, M. (2003). Putting names to faces: Successful encoding of associative memories activates the anterior hippocampal formation. *NeuroImage,* 20(2), 1400–1410. https://doi.org/10.1016/S1053-8119(03)00391-4

Sreenivasan, K. K., Vytlacil, J., & D'Esposito, M. (2014). Distributed and Dynamic Storage of Working Memory Stimulus Information in Extrastriate Cortex. *Journal of Cognitive Neuroscience*, 26(5), 1141–1153. https://doi.org/10.1162/jocn_a_00556

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, 20(11), 1643–1653. https://doi.org/10.1038/nn.4650

Stein, T., & Peelen, M. V. (2015). Content-specific expectations enhance stimulus detectability by increasing perceptual sensitivity. *Journal of Experimental Psychology: General*, 144, 1089–1104. https://doi.org/10.1037/xge0000109

Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. *NeuroImage*, 65, 69–82. https://doi.org/10.1016/j.neuroimage.2012.09.063

Stevens, C., & Bavelier, D. (2011). The role of selective attention on academic foundations: A cognitive neuroscience perspective. *Developmental Cognitive Neuroscience*, 2(Suppl 1), S30–S48. https://doi.org/10.1016/j.dcn.2011.11.001

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, 19(7), 394–405. https://doi.org/10.1016/j.tics.2015.05.004

Stokes, M., Thompson, R., Nobre, A. C., & Duncan, J. (2009). Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proceedings of the National Academy of Sciences*, 106(46), 19569–19574. https://doi.org/10.1073/pnas.0905306106

Störmer, V. S., & Alvarez, G. A. (2014). Feature-Based attention Elicits Surround Suppression in Feature Space. *Current Biology*, 24(17), 1985–1988. https://doi.org/10.1016/j.cub.2014.07.030

Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience*, 15(11), 745–756. https://doi.org/10.1038/nrn3838

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, 13(9), 403–409. https://doi.org/10.1016/j.tics.2009.06.003

Summerfield, C., & Egner, T. (2014). Attention and Decision-Making. In A. C. (Kia) Nobre & S. Kastner (Eds.), The Oxford Handbook of attention (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199675111.013.018

Summerfield, C., & Egner, T. (2016). Feature-Based attention and Feature-Based expectation. *Trends in Cognitive Sciences*, 20(6), 401–404. https://doi.org/10.1016/j.tics.2016.03.008

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., & Hirsch, J. (2006). Predictive Codes for Forthcoming Perception in the Frontal Cortex. *Science*, 314(November), 1311–1314. https://doi.org/10.1017/cbo9781139167291.033

Summerfield, C., Wyart, V., Johnen, V. M., & de Gardelle, V. (2011). Human Scalp Electroencephalography Reveals that Repetition Suppression Varies with expectation. *Frontiers in Human Neuroscience*, 5, 67. https://doi.org/10.3389/fnhum.2011.00067

Sy, J. L., Guerin, S. A., Stegman, A., & Giesbrecht, B. (2014). Accurate expectancies diminish perceptual distraction during visual search. *Frontiers in Human Neuroscience*, 8(MAY), 1–18. https://doi.org/10.3389/fnhum.2014.00334

Tang, M. F., Smout, C. A., Arabzadeh, E., & Mattingley, J. B. (2018). Prediction error and repetition suppression have distinct effects on neural representations of visual information. *eLife*, 7. Scopus. https://doi.org/10.7554/eLife.33123

Tenssay, F., & Wang, H. (2019). Analysis of EEG Signals during Visual Processing: An ERP Study. 2019 IEEE International Conference on Signal Processing, *Communications and Computing* (ICSPCC), 1–5. https://doi.org/10.1109/ICSPCC46631.2019.8960831

Theeuwes, J. (2013). Feature-based attention: It is all bottom-up priming. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1628), 20130055. https://doi.org/10.1098/rstb.2013.0055

Theeuwes, J., & Van Der Burg, E. (2007). The role of spatial and nonspatial information in visual selection. *Journal of Experimental Psychology: Human Perception and Performance*, 33(6), 1335–1351. https://doi.org/10.1037/0096-1523.33.6.1335

Thielen, J., Bosch, S. E., van Leeuwen, T. M., van Gerven, M. A. J., & van Lier, R. (2019). Evidence for confounding eye movements under attempted fixation and active viewing in cognitive neuroscience. *Scientific Reports*, 9, 17456. https://doi.org/10.1038/s41598-019-54018-z

Thomas, D. G., Grice, J. T., Najm-Briscoe, R. G., & Williams-Miller, J. (2004). The Influence of Unequal Numbers of Trials on Comparisons of Average Event-Related Potentials. *Developmental Neuropsychology*, 26(3), 753–774.

Trapp, S., Lepsien, J., Kotz, S. A., & Bar, M. (2016). Prior probability modulates anticipatory activity in category-specific areas. *Cognitive, Affective and Behavioral Neuroscience*, 16(1), 135–144. https://doi.org/10.3758/s13415-015-0373-4

Treue, S., & Maunsell, J. H. R. (1999). Effects of attention on the Processing of Motion in Macaque Middle Temporal and Medial Superior Temporal Visual Cortical Areas. *The Journal of Neuroscience*, 19(17), 7591–7602. https://doi.org/10.1523/JNEUROSCI.19-17-07591.1999

Trübutschek, D., Marti, S., Ojeda, A., King, J. R., Mi, Y., Tsodyks, M., & Dehaene, S. (2017). A theory of working memory without consciousness or sustained activity. *eLife*, 6, 1–29. https://doi.org/10.7554/eLife.23871

Uncapher, M. R., & Rugg, M. D. (2009). Selecting for Memory? The Influence of Selective attention on the Mnemonic Binding of Contextual Information. *Journal of Neuroscience*, 29(25), 8270–8279. https://doi.org/10.1523/JNEUROSCI.1043-09.2009

van Driel, J., Gunseli, E., Meeter, M., & Olivers, C. N. L. (2017). Local and interregional alpha EEG dynamics dissociate between memory for search and memory for recognition. *NeuroImage*, 149(October 2016), 114–128. https://doi.org/10.1016/j.neuroimage.2017.01.031

Van Ede, F., De Lange, F. P., & Maris, E. (2012). Attentional Cues Affect Accuracy and Reaction Time via Different Cognitive and Neural Processes. *Journal of Neuroscience*, 32(30), 10408–10412. https://doi.org/10.1523/JNEUROSCI.1337-12.2012

Varoquaux, G. (2018). Cross-validation failure: Small sample sizes lead to large error bars. *NeuroImage*, 180(April 2017), 68–77. https://doi.org/10.1016/j.neuroimage.2017.06.061

Vaziri-Pashkam, M., & Xu, Y. (2017). Goal-Directed Visual Processing Differentially Impacts Human Ventral and Dorsal Visual Representations. *Journal of Neuroscience*, 37(36), 8767–8782. https://doi.org/10.1523/JNEUROSCI.3392-16.2017

Verhoef, B.-E., Bohon, K. S., & Conway, B. R. (2015). Functional Architecture for Disparity in Macaque Inferior Temporal Cortex and Its Relationship to the Architecture for Faces, Color, Scenes, and Visual Field. *Journal of Neuroscience*, 35(17), 6952–6968. https://doi.org/10.1523/JNEUROSCI.5079-14.2015

Vetter, P., Smith, F. W., & Muckli, L. (2014). Decoding sound and imagery content in early visual cortex. *Current Biology*, 24(11), 1256–1262. https://doi.org/10.1016/j.cub.2014.04.020

Walsh, K. S., & McGovern, D. P. (2018). Expectation Suppression Dampens Sensory Representations of Predicted Stimuli. *Journal of Neuroscience*, 38(50), 10592–10594. https://doi.org/10.1523/JNEUROSCI.2133-18.2018

Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016). Reliability of dissimilarity measures for multi-voxel pattern analysis. *NeuroImage*, 137, 188–200. https://doi.org/10.1016/j.neuroimage.2015.12.012

Wandell, B. A., & Winawer, J. (2015). Computational neuroimaging and population receptive fields. *Trends in Cognitive Sciences*, 19(6), 349–357. https://doi.org/10.1016/j.tics.2015.03.009

Wen, T., Mitchell, D. J., & Duncan, J. (2018). Response of the multiple-demand network during simple stimulus discriminations. *Neuroimage*, 177, 79–87. https://doi.org/10.1016/j.neuroimage.2018.05.019

Westfall, J. (2016). PANGEA: Power ANalysis for GEneral Anova designs.

Willems, R. M., Frank, S. L., Nijhof, A. D., Hagoort, P., & van den Bosch, A. (2016). Prediction During Natural Language Comprehension. *Cerebral Cortex*, 26(6), 2506–2516. https://doi.org/10.1093/cercor/bhv075

Wimber, M., Alink, A., Charest, I., Kriegeskorte, N., & Anderson, M. C. (2015). Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nature Neuroscience*, 18(4), 582–589. https://doi.org/10.1038/nn.3973

Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*, 44(12), 1411–1426. https://doi.org/10.1016/j.visres.2003.11.024

Womelsdorf, T., Fries, P., Mitra, P. P., & Desimone, R. (2006). Gamma-band synchronization in visual cortex predicts speed of change detection. Nature, 439(7077), 733–736. https://doi.org/10.1038/nature04258

Woolgar, A., Jackson, J., & Duncan, J. (2016). Coding of Visual, Auditory, Rule, and Response Information in the Brain: 10 Years of Multivoxel Pattern Analysis. *Journal of Cognitive Neuroscience*, 28(10), 1433–1454. https://doi.org/10.1162/jocn_a_00981

Woolgar, A., Williams, M. A., & Rich, A. N. (2015). Attention enhances multi-voxel representation of novel objects in frontal, parietal and visual cortices. *NeuroImage*, 109, 429–437. https://doi.org/10.1016/j.neuroimage.2014.12.083

Wyart, V., Nobre, A. C., & Summerfield, C. (2012). Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proceedings of the National Academy of Sciences*, 109(9), 3593–3598. https://doi.org/10.1073/pnas.1120118109

Yon, D., Gilbert, S. J., de Lange, F. P., & Press, C. (2018). Action sharpens sensory representations of expected outcomes. *Nature Communications*, 9(1), 1–8. https://doi.org/10.1038/s41467-018-06752-7

Yon, D., Thomas, E. R., Gilbert, S. J., de Lange, F. P., Kok, P., & Press, C. (2023). Stubborn Predictions in Primary Visual Cortex. *Journal of Cognitive Neuroscience*, 35(7), 1133–1143. https://doi.org/10.1162/jocn_a_01997

Zuanazzi, A., & Noppeney, U. (2019). Distinct Neural Mechanisms of Spatial attention and expectation Guide Perceptual Inference in a Multisensory World. *Journal of Neuroscience*, 39(12), 2301–2312. https://doi.org/10.1523/JNEUROSCI.2873-18.2019