



Universidad Nacional Autónoma de México
Maestría en Ciencias (Neurobiología)
Instituto de Neurobiología
Departamento de Neurobiología Conductual y Cognitiva

Estudio de los correlatos neurales de la percepción
emocional por análisis de patrones en multitud de
voxeles

Que para optar por el grado de:
Maestro en Ciencias
Isaac David Reyes González

Tutor principal:
Dr. Fernando Alejandro Barrios Álvarez
Instituto de Neurobiología, UNAM

Comité tutor:
Dr. Eduardo Adrián Garza Villarreal
Instituto de Neurobiología, UNAM
Dr. Luis Concha Loyola
Instituto de Neurobiología, UNAM

México, enero de 2020

Abstract

Metanalyses in affective neuroscience have failed to agree on the specific neural substrates of distinct emotional experiences. Task-related fMRI has long relied on multiple-regression analysis of single-voxel signals (notably GLM), which assumes that activity at some region will correlate with experimental variables, but this needs not be the case: cognitive states might emerge from the *joint* activity probability distribution of many elements. This has led many to believe that emotion decoding could benefit from the more sensitive machine-learning methods of multi-voxel pattern analysis (MVPA). However, so far these have only been used to either explore one region at a time, or with no intention to derive group-level activation imaging when used brain-widely. This work explores the extent to which MVPA can overcome these limits as an actual substitute for GLM, and applies it to the problem of narrowing down the neural substrates of emotion *perception*.

Two similar experiments were performed, each one using a separate validated stimuli dataset on a different control group of people, who underwent their respective emotion perception task on a 3 T MR scanner. Subjects were presented with neutral, happy, sad and angry (second experiment only) faces in equal amounts at random. For the first experiment, short blocks of interest were interspersed with a variable amount of neutral control blocks, so as to avoid both subject habituation and test-partition contamination by temporal correlation. The second experiment addressed complexity and confounding concerns in the former, and amassed 800 % more training data.

After image preprocessing, linear SVM classifiers (one per subject) were trained for all non-trivial block combinations. Predictive power was estimated using cross-validation, then compared to the null hypothesis distribution estimated using the Monte-Carlo method (5000 training-partition label permutations). Instead of assuming an optimal labeling delay after stimulus presentation, the timeseries of classification accuracy were measured for all possible delays given fMRI's temporal resolution; and group-level hypothesis testing was conducted considering the single delay that maximized average

accuracy for a particular class contrast (first and second experiment), and also the best delay per subject (first experiment only, corrected for multiple comparisons).

Average classification timeseries' morphology agrees with theoretical expectations about the hemodynamic response, although little to no evidence of successful decoding was found for the first database; even after some attempts at dimensionality reduction (top 1 % of gray-matter voxels according to their F-test score), which pushed 3-way classification performance to 35 % accuracy (delays: 2.2 s and 5.2 s). When allowing different delays per subject, accuracy surpassed the 50 % line to great statistical effect ($D_{Cohen} = 3,7$), but was likely to be a coincidence ($p_{Bonferroni} = ,22$).

Depending on emotion combination, average classification timeseries for the second database saw most maxima occur 4 s after stimulus onset, and never immediately or too late. Performance varied from random guessing to statistical significance; yet this variability followed a pattern: only all combinations including *happiness* showed evidence of successful decoding. E.g. happy vs angry ($p = ,04$, $D_{Cohen} = 2,2$), happy vs sad ($p = ,07$, $D_{Cohen} = 1,9$), happy vs sad vs neutral ($p = ,09$, $D_{Cohen} = 1,7$). No dimensionality reduction was performed other than gray-matter masking. Moreover, statistical mapping ($z_value > 2,3$) of normalized SVM weights for successful contrasts consistently threw structures implicated in facial and emotional processing: anterior cingulate, orbitofrontal, insular and fusiform cortices; amygdalas and ventral striatum. Meanwhile, classical analysis with GLM was unable to find any significant activation cluster.

These results indicate that standard fMRI techniques can record differential information pertaining to brain activity associated with certain emotions (notably happiness), so a multivariate algorithm like SVM could learn to decode such slippery states above chance levels, whereas the univariate algorithm couldn't; statistical power being equal. The anatomical consistency of discriminating features across subjects and contrasts despite high-dimensionality and a small subject sample, as well as their correspondence with the wider literature, suggests MVPA may be a viable tool for functional neuroanatomical mapping and not just prediction of psychological states.

Índice general

Abstract	i
1. Introducción	1
2. Antecedentes	8
2.1. Bases neuronales de las emociones	8
2.2. Análisis de patrones multivoxel	10
3. Justificación	15
4. Hipótesis	23
4.1. Preguntas de investigación	23
4.2. Hipótesis	23
4.3. Predicciones	23
5. Objetivos	24
5.1. General	24
5.2. Específicos	24
6. Muestra, Materiales y Métodos	25
6.1. Participantes	25
6.2. Estímulos y diseño experimental	26
6.3. Adquisición de imágenes	30
6.4. Métodos de análisis	31
7. Resultados	32
8. Discusión	33
9. Conclusiones	34

Apéndice 1: código fuente	35
Referencias	37

Capítulo 1

Introducción

Las emociones son reacciones corporales estereotipadas ante estímulos o eventos, internos o externos, que el organismo evalúa como relevantes en términos de las muchas facetas del éxito evolutivo (Schacter et al. 2011; LeDoux, Damasio 2013). Por ejemplo, el miedo de un gato espeluznado ante un perro — como cualquier instancia de esta emoción — es una respuesta adaptativa forjada en el fenotipo de las especies o en su conducta aprendida para proceder con cautela durante la percepción de peligro. La reacción de una emoción abarca tanto componentes psicológicos como meramente fisiológicos (Schacter et al. 2011; LeDoux, Damasio 2013). En el caso del desafortunado gato se tendrían, por un lado, la relocación súbita de la atención, la sensación misma del miedo como una suerte de compresión de abajo hacia arriba de las entrañas; y por otro, la sobreapertura de los párpados, hiperventilación y aceleración del ritmo cardíaco, liberación de hormonas estresantes, piloerección y quizás conductas de congelamiento por parte de la musculatura estriada.

La familia de las emociones es más numerosa que el miedo, si bien las relaciones precisas entre sus miembros son un punto de debate. Por mencionar algunas de las más populares, fenómenos como la alegría, la tristeza, el enojo, el asco y la vergüenza también forman parte de esta familia (Ekman, Davidson 1994). Cada una ameritaría su propia descripción en tanto que fenómeno psicofisiológico, así como su posible explicación funcional.

Existe otro sentido, entretelado al anterior, en el que se usa la palabra “emoción”: el sentido de “sentimiento” (feeling). Desde esta segunda acepción, la emoción denota la sensación o experiencia consciente¹ que suele acompañar

¹Por “consciencia” me referiré específicamente al significado amplio y rudimentario de consciencia fenoménica: la cualidad de los objetos que sienten; en la línea de autores como Christof Koch, David Chalmers, Giulio Tononi y Ned Block.

a los exabruptos fisiológicos anteriormente mencionados (Tsuchiya, Adolphs 2007; Schacter et al. 2011; LeDoux, Damasio 2013). Dos características parecen ser necesarias y suficientes para distinguir la fenomenología de las emociones respecto de otros estados mentales (Schacter et al. 2011): la primera es que hacen referencia a su respectivo estado somático;² y la segunda, no sorprendentemente, es que acarrean un valor hedónico. Es decir, las emociones parecen siempre ser agradables o desagradables, sin lugar a una tercera opción.³

Las emociones son un fenómeno *muy* básico y bien conservado a través de la filogenia, tanto así que algunos estudiosos sugieren que los rudimentos de las manifestaciones conductuales de las emociones (esencialmente conductas de apertura-acercamiento o rechazo) son trazables a respuestas análogas en todo el árbol de la vida (Damasio, Carvalho 2013); aunque el caso más claro por mucho es el de animales con un sistema nervioso, que es lo que sustenta y coordina sus funciones emocionales. Entonces, como función básica del sistema nervioso, resulta interesante en su propio derecho entender ese sustrato. Y aunque nadie sano duda poder distinguir la alegría de la tristeza (ya sean los propios o ajenos), descubrir su naturaleza en actividad nerviosa ha probado ser especialmente difícil.

La problemática es patente a partir del abanico de teorías psicológicas que compiten por caracterizarlas en términos conceptuales, materiales y evolutivos. Por un lado se encuentran los modelos que postulan la existencia de emociones “básicas” o atómicas, que pueden variar en intensidad y combinarse para producir emociones más *sui géneris*. Otros modelos las sitúan a todas en un espacio continuo de componentes afectivos más simples (como valencia hedonística y nivel de excitación corporal). En tercer lugar encontramos las teorías de “appraisals”, para las cuales la categorización emocional no se conforma a ningún espacio de dimensión finita, sino que surge de evaluaciones situacionales más finas conforme al contexto (ejemplos: soledad, decepción o culpabilidad en lugar de llana tristeza). Otros más terminarían por argüir

²A diferencia de, dígame, el exquisito olor de las flores; que se corresponde a un objeto externo. Damasio (1996) utiliza el término “somático” en lugar de “corporal” para comunicar la suficiencia de *representaciones* o simulaciones mentales del cuerpo.

³A diferencia de, por ejemplo, la experiencia libre de placer o desagrado de un color vívido, o información propioceptiva sobre la posición de las extremidades.

que no hay modelos unificadores puesto que se trata de fenómenos corporales muy dispares, y los intentos por agruparlos son poco más que constructos culturales (ver Kragel, LaBar 2016 para una revisión comparativa).

Sin importar cuál sea su organización, sabemos que las predisposiciones en la dinámica afectiva del individuo son una importante marca de la personalidad (Schacter et al. 2011), que como cualquier mecanismo corporal está sujeto a descomposturas. Podría parecer extraño hablar acerca de estados psicofisiológicos de ocurrencia natural y frecuente en relación a la salud; sin embargo su matrimonio no puede ser pasado por alto. En primer lugar, las emociones mismas pueden ser sintomáticas de la salud o la enfermedad (como es el caso del asco), ya que juegan un papel clave en la representación neuronal del estado organísmico global (Damasio 1999). Pero más directamente, se reconoce que afecciones a los sistemas de la emoción producen trastornos o psicopatologías (Berrios 1993; Millon, Davis 1996); independientemente de la selección de entidades clínicas y diagnósticas del momento, y del incipiente conocimiento científico de sus etiologías o causas concretas. Se pueden mencionar, a manera de ilustraciones, diversas formas de depresión extendida (trastorno depresivo mayor, distimia), anhedonia, fobias y desórdenes de la ansiedad, así como el trastorno limítrofe o “emocionalmente inestable” de la personalidad.⁴

Aunque de menor prevalencia que las disregulaciones psiquiátricas anteriores, funciones emocionales completas pueden llegar a perderse mediante lesiones, síndromes genéticos autosómicos recesivos, patógenos y otras formas de deterioros neurológicos focalizados a estructuras del sistema emocional. Entre estas estructuras cerebrales diversas encontramos núcleos mesencefálicos, diencefálicos, telencefálicos subcorticales en ganglios basales y en la amígdala,⁵ así como partes de la neocorteza de los mamíferos;⁶ resaltando la

⁴Para listas y descripciones más exhaustivas de uso corriente entre psiquiatras y psicólogos clínicos, se pueden consultar las varias entradas sobre desordenes depresivos y de ansiedad en el Manual diagnóstico y estadístico de los trastornos mentales (American Psychiatric Association 2013), o las entradas sobre desordenes afectivos y relacionados con el estrés en el capítulo 10 de la Clasificación internacional de enfermedades (World Health Organization 2004).

⁵Lo que ocasiona el síndrome de Klüver-Bucy (Marlowe et al. 1975) o la enfermedad de Urbach-Wiethe (Hurlemann et al. 2007).

⁶La demencia frontotemporal en su variante conductual es un caso de síndrome neurológico de origen cortical con fuertes repercusiones emocionales (Rohrer et al. 2011).

segregación y complejidad de los sistemas de la emoción. Cabe mencionar que la investigación biomédica en torno a problemas de salud emocional no se limita al *localizacionismo*. Existe un fuerte interés por encontrar explicaciones y biomarcadores clínicos modelando al cerebro como una red de componentes en interacción (Rubinov, Sporns 2010). Por ejemplo, se ha hipotetizado que el síndrome de Capgras resulta de una pérdida de comunicación entre el sistema de reconocimiento de caras en el lóbulo temporal y el sistema emocional, no obstante que ambos siguen operando bien por separado (Ellis et al. 1997; Ramachandran 1998).

Lejos de ser un mecanismo aislado — ¡y en contra del sentido común! —, la cognición más superior y racional es y *debe ser* en cierta medida influida y modulada por emociones. En estado fisiológico, la gente usa sus corazonadas emocionales como información para nutrir sus creencias sobre la probabilidad de éxito en una actividad (Isen, Patrick 1983). ¿Por qué sería útil y heurístico este sesgo? La respuesta es simple: el mundo causa en gran medida nuestros estados afectivos; por lo tanto los estados afectivos pueden proveer información sucinta — vaya, de calidad ejecutiva — acerca del mundo (Schwarz, Clore 2003). Como pieza en el rompecabezas de las funciones neurales, la emoción provee cimientos para algo tan crucial y etéreo como la motivación (Schwarz 1990; Schacter et al. 2011); y por transitividad, sobre éste también descansa la planificación, sin mencionar aspectos de la toma adecuada de decisiones (Schwarz 1990; Bechara et al. 2000), ambos característicos de la inteligencia. ¿Cuál es la *causa final* de que el granjero ahorre sus ingresos para comprar un nuevo tractor? ¿Es transformar la maleza, es aumentar la producción, obtener incluso más ingresos, es comprar otros bienes, o es la búsqueda de un bienestar físico y psicológico para él y su familia?

También aquí la patología ofrece una oportunidad para poner de manifiesto la relevancia cognitiva de la emocionalidad. La paciente S.M., quien perdió ambas amígdalas debido a la enfermedad de Urbach-Wiethe, mantiene un desempeño ileso en pruebas comunes de conocimientos, reconocimiento, inteligencia fluida y memoria. ¿Qué la lleva entonces a apostar y perder todo su dinero en simulaciones,⁷ a mostrar dificultades decidiendo sobre la prioridad

⁷Como la prueba neuropsicológica del juego de azar de Iowa (Iowa Gambling Task) (Bechara et al. 2001) y la prueba por desliz (Faux Pas) (ver Torralva et al. 2009).

de las opciones que se le presentan en su vida cotidiana, y a mostrarse transgresora al punto de la ineptitud interpersonal? El problema no está en su habilidad para razonar, sino en su incapacidad para sentir. La ausencia de instintos como el miedo se traduce en una pésima adopción de estrategias en situaciones emocionalmente cargadas (Damasio 1996; Bechara et al. 2000).

Más allá del individuo, existe especial significancia sobre lo que las emociones logran para el bienestar social. Este texto inició por plantear el concepto de emoción como una reacción corporal adaptativa para la supervivencia más inmediata y prosaica, sin embargo esta visión es miope. Es sumamente probable que entre el complejo expresivo y postural de una emoción también encontremos características seleccionadas para la comunicación, y no necesariamente para predisponer al organismo a sortear la situación o aprender de ella (Russell, Dols 1997; Tracy et al. 2015). A continuación se exponen distintos niveles en los cuales la emoción opera como facultad cognitiva social, desde la selección de grupo y sexual hasta las sofisticadas mentalización y empatía.

En etología (el estudio del comportamiento animal) se habla de que una señal del organismo emisor al receptor es efectiva si posee la propiedad de ser ventajosa para ambas partes, en promedio (Donath 2007; Laidre, Johnstone 2013). Por ejemplo, la aparatosa aposematosis de una indefensa rana venenosa, o las costosas estridulaciones a triple *forte* del grillo macho son más que un regalo para disuadir predadores o atraer parejas, respectivamente. Claramente, la rana se beneficia de no ser comida aunque como bocadillo sea letal, así como el beneficio es compartido entre ambos sexos de grillo. Lo mismo estaría pasando con el llanto de un niño, o con la excesiva hilaridad con la que actuaba su madre cuando se interesó por primera vez en el padre: el llanto también existe para beneficio de la madre, y la jocosidad con la que casi involuntariamente comenzó a seducir al padre — en su momento — fue más que una autoseñalización de estar en un ambiente agradable y confiable.

Más aun, si es difícil fingir o suprimir estas señales (en proporción a la flexibilidad cognitiva de la especie), entonces mayores son los indicios de su efectividad a lo largo de la historia de la evolución. Mentir o prevaricar resta al valor que receptores y emisores honestos obtienen de dichas señales (Do-

nath 2007; Tracy et al. 2015); por lo que resulta sumamente interesante para etólogos y antropólogos notar que, incluso en las especies más inteligentes, la señalización emocional es difícil de fingir (Damasio 1999; Ekman 2003; Porter, Ten Brinke 2008). En efecto, la sociedad humana reconoce a los buenos actores teatrales como ejecutores de proezas dignas de premios.

Es en nada más y nada menos que primates donde la percepción⁸ de expresiones emocionales parece haber asido circuitos sumamente especializados, específicamente usando la expresión facial como medio (Tsao et al. 2006). Pocos estímulos son tan informativos como los rostros. Basta una fracción de segundo para que la cara de una persona nos cuente algo acerca de su sexo, edad, etnia; incluso de su atención y por supuesto, su estado anímico (Tsao, Livingstone 2008; Haist, Anzures 2017). Quizás ningún otro induce tantas ilusiones pareidólicas, y la sofisticación de su reconocimiento — que aparece pocas horas después de nacer y alcanza la cuasi-madurez naturalmente antes del primer año (Haist, Anzures 2017) — solamente es equiparable a habilidades aprendidas arduamente como la lectura.

Se sabe al menos desde tiempos de Hughlings Jackson y Charcot que el cerebro humano puede perder selectivamente la capacidad para identificar rostros. Este desorden neurocognitivo, conocido como ceguera o agnosia de caras (prosopagnosia), depende de regiones corticales especializadas en la vía visual ventral (la vía del reconocimiento de objetos); específicamente de cierta área en los giros fusiformes (“Fusiform Face Area”) (Haxby et al. 2000; Said et al. 2010).⁹ Mucho antes de la aparición del lenguaje *bona fide* humano, nuestros ancestros tuvieron que leer estas expresiones e interpretarlas como parte de un repertorio de señalización para la vida en pequeñas sociedades de tropa (Fessler, Gervais 2010; Fridlund 2014). Acordemente, las mismas técnicas modernas de neuroimagen muestran evidencia de áreas análogas a FFA en macacos (Tsao et al. 2006), cuyo ancestro en común con humanos

⁸Percepción es el conjunto de procesos de interfaz mediante los cuales un sistema cognitivo adquiere información externa y la interpreta, volviéndola útil para procesos de orden superior como la categorización y conceptualización, la acción y el control ejecutivo.

⁹Bien explican Haxby y coautores (*ibidem*) y Haist, Anzures (2017) que la actividad del surco temporal superior posterior (pSTS) también se ha asociado al procesamiento de los aspectos faciales dinámicos; y tanto FFA como pSTS se alimentan del lóbulo occipital lateral inferior ipsilateral. Por lo tanto, se habla de toda una red de procesamiento de caras.

data de hace más de 25 millones de años (Lewis 2013). Para rematar, exploraciones más invasivas mediante arreglos de microelectrodos han revelado que las neuronas en estos “parches” del macaco responden exclusivamente a rostros (Tsao et al. 2006). Es decir, no solo el agregado de actividad LFP¹⁰ sigue un código para la detección de rostros; el campo receptivo de cada célula individual de alguna manera ya lo está haciendo.

Este sobrevuelo por los extensos territorios de las emociones no podría finalizar sin rozar las cumbres más elevadas del antropocentrismo. Después de la supervivencia, la salud, la cognición y la visita al zoológico... la buena percepción de emociones termina por alimentar los mecanismos de algo tan sofisticado como la teoría de la mente¹¹ (Mitchell, Phillips 2015) y la empatía (Shamay-Tsoory et al. 2009; Zaki, Ochsner 2012). Al ser éstas una flagrante muestra de los estados mentales más viscerales, todo el tiempo estamos leyendo las marcas de su expresión en el otro, retroalimentando nuestros propios aparatos cognitivos para poder lidiar con ellos, o bien, empatizar y sentir con ellos. Quizás una explicación de este tipo, un elaborado refriego de reflejos de antaño, subyazca controvertidos fenómenos (para las ciencias biológicas) como la adopción y la compasión que otras especies nos merecen. Debido a esto, más allá de ser ajenas al humano y lo humano, las emociones son parte y arte de los complejos razonamientos que acarrea la vida en sociedad.

¹⁰Local Field Potential: potencial de campo local.

¹¹Teoría de la mente: atribuir estados mentales a otros. La creencia de que existen otras mentes; en oposición al solipsismo. No debe confundirse con las teorías de lo que la mente es.

Capítulo 2

Antecedentes

2.1. Bases neuronales de las emociones

Las investigaciones del procesamiento emocional hasta antes de la neuroimagen funcional lograron implicar varias regiones corticales y subcorticales; algunas de ellas más bien primitivas y bien conservadas. A grandes rasgos, los impulsos tempranos que entran al tálamo son copiados a corteza sensorial pero también a la amígdala. Ésta última contiene núcleos especializados para evaluación de carga emocional tanto innata como aprendida. Una larga historia de investigación en condicionamiento ha mostrado que la amígdala debe estar presente para aprender asociaciones implícitas (es decir, procedimentales) entre estímulos que generan miedo y estímulos incondicionados. El núcleo central de la amígdala proyecta a lugares como la región gris central del tallo cerebral e hipotálamo, que a su vez puede desencadenar reacciones automáticas y actividad humoral. También hay áreas corticales implicadas en la percepción de emociones complejas: la corteza cingulada anterior (ACC), la ínsula y la corteza prefrontal ventromedial (vmPFC). Como ejemplo de las consecuencias que estas estructuras tienen en percepción emocional, daño a la corteza prefrontal resulta en discapacidades sociopáticas (LeDoux, Damasio 2013).

No obstante, la localización y codificación de emociones tanto en el sistema nervioso central como en el periférico permanece un problema abierto (Kragel, LaBar 2016; Celeghin et al. 2017). Una pregunta fundamental es si existen patrones en actividad neuronal suficientemente específicos que correspondan con concepciones vernáculas y académicas de las emociones. Además, ¿qué tan generalizables son esos correlatos entre individuos y especies? Esta última cuestión tiene acaloradas implicaciones para la comprensión de la

aparición de estados afectivos, ya sea en términos de evolución genética o cultural.

Los metanálisis y revisiones de la literatura hasta antes de la aplicación de detección de patrones multivariados no muestran una tendencia en la respuesta a las preguntas anteriores. Phan et al. (2002) analizaron los mapas de activación de 55 estudios de PET¹ y fMRI², y encontraron evidencia parcial para la existencia de correlatos neuroanatómicos consistentes a la ocurrencia de distintos tipos de emociones básicas. Murphy et al. (2003) incrementaron el tamaño de muestra a 106 estudios y mantuvieron la conclusión general respecto del metanálisis anterior, aunque los correlatos no fueron suficientemente similares. La relativa debilidad de los resultados ha sido usada como evidencia de hipótesis alternativas; a saber, teorías dimensionales y de actos conceptuales (emociones como construcciones de afectos, Barrett 2006). Barrett, Wager (2006) examinaron 161 estudios y encontraron correlatos exclusivos (mas no específicos) para miedo, tristeza y asco. La felicidad y el enojo no arrojaron correlatos ni consistentes ni específicos. Vytal, Hamann (2010) utilizaron un análisis con mayor sensibilidad espacial para contrastar cinco emociones básicas provenientes de 83 estudios. Concluyeron que aunque existen *clusters* característicos de activación para cada una de ellas, no es evidencia suficiente para descartar otros modelos de representación emocional en el cerebro. Lindquist et al. (2012) analizaron 91 estudios tanto de experiencia de emociones como de percepción emocional. Definieron evidencia a favor de emociones básicas como actividad consistente en áreas específicas selectivas a una única categoría. Se encontró consistencia y selectividad para las cinco categorías bajo consideración, mas no especificidad funcional. Es decir, las escurridizas emociones básicas tienen correlatos distinguibles entre sí, aunque las mismas regiones cerebrales están vinculadas a otras funciones psicológicas extraemocionales.

El consenso disponible es que la búsqueda de estructuras específicas a un tipo de emoción está superada (Murphy et al. 2003; Kober et al. 2008; Lindquist et al. 2012). Estructuras cerebrales que parecían tener una correspondencia

¹Positron Emission Tomography: tomografía por emisión de positrones

²Functional Magnetic Resonance Imaging: imagenología de resonancia magnética funcional.

uno-a-uno con alguna emoción básica, como la amígdala (miedo) y la ínsula (asco) (Calder et al. 2001), ahora son acreedoras a funciones más elusivas o generales (Sander et al. 2003). Si la función emocional tiene biomarcadores reproducibles, su decodificación dependerá de la actividad conjunta de redes funcionales distribuidas y traslapadas (Hamann 2012; Lindquist, Barrett 2012; Kragel, LaBar 2014; Celeghin et al. 2017; aunque no se han realizado tantos estudios electrofisiológicos, la revisión de Guillory, Bujarski 2014 concuerda con esta conclusión).

2.2. Análisis de patrones multivoxel

Dado que los métodos clásicos de análisis de patrones diferenciales de activación en fMRI y PET están agotando su utilidad (notablemente el algoritmo GLM,³ basado en modelos de regresión lineal múltiple para voxeles individuales), el campo de neuroimagen funcional orientada a tareas comienza a recurrir a análisis verdaderamente multivariados para dilucidar las preguntas de interés. Estos diversos métodos, también agrupados bajo el nombre de “análisis de patrones multivoxel” en la literatura neurocientífica (MVPA por sus siglas en inglés), comparten la característica de modelar la codificación del estado mental, cognitivo o conductual como una función de muchas dimensiones; que en el caso de fMRI toman la forma de voxeles distantes en espacio o tiempo (Norman et al. 2006; Huettel et al. 2009; Mahmoudi et al. 2012; Lewis-Peacock, Norman 2013). Al considerarlas holísticamente, MVPA es capaz de extraer igual o mayor información de las burdas imágenes funcionales en la mayoría de los casos (ver Jimura, Poldrack 2012 para contraejemplos), a expensas de complicar potencialmente el modelo y su cómputo. La adopción de este enfoque en estudios con fMRI ha llevado a descubrimientos novedosos en casi todas las ramas de la neurociencia cognitiva (por ejemplo Haynes, Rees 2005; Polyn et al. 2005; Soon et al. 2008; Rissman, Wagner 2012; véanse Huettel et al. 2009; Tong, Pratte 2012 para más).

³General Linear Model; modelo lineal general. En el contexto de neuroimagen funcional, consúltense Friston et al. (1994) y Penny et al. (2011).

Según Lewis-Peacock, Norman (2013), el término MVPA es usado para referirse a dos estrategias distintas de análisis de patrones multivoxel; ambas provenientes del campo de estudio en el que se intersecan la estadística inferencial multivariada y la inteligencia artificial, un área conocida como “aprendizaje de máquinas” o “aprendizaje automático” (“machine learning” en inglés):

- Clasificación automática empleando algoritmos de aprendizaje supervisado.
- Cuantificación de similitud entre patrones, según alguna métrica de distancia.

Mientras que los primeros decodifican estados preestablecidos a partir de la actividad registrada, la segunda línea de análisis se asemeja a la teoría de los modelos de aprendizaje automático no supervisado; y puede entre otras cosas usarse para discernir si dos zonas del sistema nervioso codifican información similar (Lewis-Peacock, Norman 2013).

Para 2012 ya se habían realizado un puñado de estudios de decodificación emocional con análisis multivariado, y su potencial era reconocido por la comunidad en trabajos revisorios (como Hamann 2012; o Schirmer, Adolphs 2017 en lo que respecta a percepción emocional; ver Kragel, LaBar 2014; y Kragel, LaBar 2016 para revisiones abocadas al uso de MVPA en materia de emociones). También es necesario recalcar que aunque muchas de las aplicaciones de MVPA en neurociencia afectiva han usado percepción de estímulos emocionalmente cargados, ésta no es exactamente análoga a la experiencia en primera persona de una emoción (notable también a partir de actividad cerebral (Wager et al. 2008)). La percepción de emociones en estímulos sociales además puede involucrar procesos de evaluación, atribución de estados mentales e intenciones; a la vez que puede estar exenta de emoción *per se* (debido a la modulación empática y sus trastornos) (Peelen et al. 2010). En lo subsecuente se hará la distinción, poniendo énfasis en los trabajos relativos a *clasificación* de estados de *percepción* emocional.

Pessoa, Padmala (2007) demostraron la factibilidad de usar MVPA para predecir la respuesta conductual en una tarea de fMRI sobre percepción de

caras asustadas, presentadas casi subliminalmente, en contraste con otros tipos de caras. Se basó en un algoritmo de entrenamiento SVM sin kernel, cuyo desempeño aumentaba conforme las variables dependientes en el modelo eran adicionadas con nuevas regiones de interés (ROIs por sus siglas en inglés); hipotetizando que los varios aspectos de la actividad emocional (discriminación del estímulo, producción de respuestas, etc.) contienen información útil para recuperar el estado perceptual. Ethofer et al. (2009) utilizaron estímulos auditivos con carga paralingüística emocional. Su modelo SVM fue capaz de clasificar entre cuatro emociones básicas y un estado aemocional por encima del azar, mientras que un análisis univariado no encontró diferencias en las mismas señales BOLD⁴, provenientes de corteza auditiva (ver Kotz et al. (2012) para un estudio similar no restringido a corteza auditiva). Said et al. (2010) utilizaron clasificación multivariada mediante el método de regresión logística para demostrar que la actividad de ciertas regiones del surco temporal superior (STS, previamente vinculadas a detección de expresiones faciales) predice la categoría emocional de los rostros. Los mismos datos bajo análisis univariado fueron incapaces de repetir la hazaña. Peelen et al. (2010) probaron estímulos de cinco categorías emocionales en modalidades distintas (audición, visión de gesticulaciones y de lenguaje corporal), y realizaron una búsqueda de regiones codificantes en todo el encéfalo mediante el método de “searchlight”. Sorprendentemente, el uso de corteza prefrontal medial (mPFC), o bien del STS, resultó capaz de hacer la distinción entre emociones, independientemente de la modalidad sensorial e intensidad del estímulo. En cambio, la aplicación de GLM sólo replicó este hallazgo supramodal en el caso del miedo. Wegrzyn et al. (2015) exploraron varias ROIs⁵ asociadas a procesamiento de rostros para comparar la magnitud en la que contienen información emocional. Todas resultaron aptas, especialmente el STS y la bien conocida “área fusiforme de caras” (FFA) en el giro fusiforme de los lóbulos temporales (especialmente el derecho).

El uso de MVPA en predicción de percepción emocional también ha alcanzado a otros modelos de las emociones y a otras especies. Skerry, Saxe (2015) utilizaron SVM para probar que las emociones percibidas, entendidas como

⁴Blood-Oxygen-Level Dependent: [Señal] dependiente del nivel de oxígeno en la sangre; la forma de contraste más común en estudios de fMRI.

⁵Region of Interest: región de interés.

“appraisals” o evaluaciones sociales finas y dependientes del contexto, pueden decodificarse por encima del azar a partir de redes cerebrales asociadas a la facultad de teoría de la mente. Hernández-Pérez et al. (2018) (prepublicación) entrenaron perros para permanecer quietos y despiertos en un resonador magnético mientras observaban rostros humanos felices o inexpresivos. En un segundo experimento emplearon más emociones y SVM para proveer evidencia de que los correlatos encontrados en la corteza temporal derecha del perro contienen información selectiva a la felicidad de los humanos.

En lo que respecta a experimentación en primera persona de emociones, los métodos de clasificación multivariada han sido capaces de caracterizar con precisión los dispersos correlatos espaciales de la valencia hedonística y el nivel de excitación (arousal) postulados por hipótesis dimensionales (Rolls et al. 2009; Baucom et al. 2012; Chikazoe et al. 2014; Shinkareva et al. 2014; Chang et al. 2015). Además hay evidencia de cierta generalidad intersujeto, pues existen patrones compartidos que permiten su clasificación a niveles de exactitud no triviales. De la misma forma se han encontrado resultados prometedores categorizando experiencias de emociones discretas (así se quieran conceptualizar como básicas o constructos), en comparación con los resultados del cúmulo de literatura con análisis univariado (Sitaram et al. 2011; Kassam et al. 2013; Saarimäki et al. 2015; Kragel, LaBar 2015). Está por verse si el uso de MVPA será suficiente para encontrar una hipótesis vencedora.

Cuadro 2.1: Muestra representativa de estudios experimentales sobre los correlatos neurales de la percepción de emociones que han empleado MVPA.

Estudio	Modalidad de estímulo	Emoción de estímulo	ROI	Algoritmo	Desempeño
Pessoa, Padmala (2007)	visual	miedo	varias	SVM	78 % > 50 %
Ethofer et al. (2009)	auditiva	alegría, enojo, tristeza, alivio	corteza auditiva	SVM	33 % > 20 %
Said et al. (2010)	visual	alegría, enojo, tristeza, asco, sorpresa, miedo	surco temporal superior	regresión logística	22 % > 14 %

Estudio	Modalidad de estímulo	Emoción de estímulo	ROI	Algoritmo	Desempeño
Peelen et al. (2010)	auditiva, visual (gesticulaciones, lenguaje corporal)	alegría, enojo, tristeza, asco, miedo	encéfalo (searchlight)	no especificado	no especificado
Kotz et al. (2012)	auditiva	alegría, enojo, tristeza, sorpresa	encéfalo (searchlight)	SVM	28 % > 20 %
Skerry, Saxe (2015)	lectura, visual (caras)	21 distinciones finas contexto-dependientes (appraisals)	red de teoría de la mente, encéfalo (searchlight)	SVM	~8 % > 5 % (ROIs)
Wegrzyn et al. (2015)	visual (caras)	alegría, enojo, miedo	varias	MGPC	~32 % > 25 %
Hernández-Pérez et al. (2018)	visual (caras)	alegría	(<i>canis familiaris</i>) corteza temporal derecha, caudado. Encéfalo (searchlight)	SVM	~65 % > 50 % (ROIs)

Capítulo 3

Justificación

Al día de hoy, clínicos y neurocientíficos cognitivos siguen siendo bastante dependientes de reportes verbales y otras pistas conductuales para evaluar estados subjetivos. Pruebas neurales análogas deberían idearse y descubrirse, pero no solo por sus aplicaciones terapéuticas y éticas (Seth 2010; Tononi, Koch 2015; Michel et al. 2019). Más general y fundamentalmente, se requiere trabajo teórico y empírico extra para tender un puente explicativo entre los hechos de la subjetividad (mentales) y los modelos del mundo objetivo (como el conocimiento científico) (Chalmers 1995; Seth 2010; Oizumi et al. 2014). Resulta agravante para los estudios de la consciencia que ni siquiera se tengan bien identificados los correlatos neuroanatómicos más macroscópicos de prominentes estados fenomenológicos, como las emociones.

Hay varios sentidos en los que las emociones básicas podrían considerarse algunos de los estados organizmicos más primitivos, después de sensaciones como el dolor (Simon et al. 2006) o un estado básico de vigilia reactiva/alertamiento (que Damasio 1999 denomina “emociones de fondo”): son universales culturales relativamente diferenciables entre sí (Ekman, Davidson 1994), y están inextricablemente conectadas a la homeostasis y estados somáticos de gran relevancia para el organismo; de modo que su correcto procesamiento en el sistema nervioso tiene una función preponderante en los mecanismos de la motivación, la toma de decisiones y el control ejecutivo (Schwarz 1990; Damasio 1996; Damasio 1998; Damasio 1999; Bechara et al. 2000).

Desde el punto de vista de la percepción emocional, ayudar a entender los mecanismos en tercera persona para su detección y procesamiento contribuye a nuestra comprensión de la intersubjetividad. Así mismo, el reconocimiento de expresiones faciales es una forma primaria de comunicación no verbal (Fridlund 2014; Russell, Dols 1997; Simon et al. 2008) que involucra la fun-

ción de vastos y distribuidos circuitos neuronales, los cuales se encuentran bajo investigación activa (Fridlund 2014; Kragel, LaBar 2014; Kragel, LaBar 2015).

Pese al éxito inicial identificando y describiendo los sustratos moleculares y celulares de algunos fenómenos psicológicos y afecciones psiquiátricas; los métodos de la neurociencia a escalas pequeñas y medianas, así como el llano *localizacionismo*, fracasan frente a cantidad de preguntas de interés. Muchos fenómenos psicológicos más bien emergen (o así se piensa) de la función coordinada de componentes anatómicamente distribuidos (Gazzaniga, Ivry 2013; Oizumi et al. 2014); y la percepción de expresiones faciales no es la excepción (Kragel, LaBar 2014). Las restricciones evolutivas imponen límites en el número de especializaciones nerviosas que podrían corresponder unívocamente con el desempeño de una función, de manera que ciertos fenómenos cognitivos deben corresponder al reclutamiento diferencial y modulación temporal de recursos más primitivos en el dominio fisiológico (Gazzaniga, Ivry 2013). Más aún, las nuevas corrientes en estudios de la cognición hacen hincapié en considerar holísticamente no solo la actividad del sistema nervioso central como si se tratara de un sistema confinado; sino la corporización de la cognición como un mecanismo evolucionado para la acción y la predicción activa (Friston 2010), e incluso considerando las vivencias con el medio (ontogenia y filogenia) y las restricciones que éste llega a imponer sobre los aparatos cognitivos (Newen et al. 2018). Naturalmente, dichos enfoques han sido bienvenidos por algunos estudiosos de la emoción de manera independiente al enfoque corporizado pionero de Damasio (por ejemplo Harrison et al. 2010; Seth 2013; Slaby 2014).

Para abordar este tipo de preguntas, la resonancia magnética nuclear funcional es capaz de registrar señales correlato de los potenciales locales de campo, a la vez que provee un moderado balance de resolución espacial y temporal para diferenciar casi simultáneamente el metabolismo aeróbico en todo el encéfalo. Es una técnica idónea para estudios en los que redes funcionales asociadas a tareas primero deben ser identificadas o refinadas, allanando el terreno para técnicas de medición más finas que intenten establecer la circuitería y química sináptica (Logothetis 2008).

El actual debate alrededor de los mecanismos cognitivos de la percepción de expresiones faciales se beneficiará de evidencia adicional en materia de disociación anatomo-funcional; y lo que es más apremiante: el uso de métodos de análisis más sensibles para derivar la evidencia necesaria.

Los métodos multivariantes han demostrado tener mayor éxito extendiendo la identificación de interacciones complejas de actividad cerebral que no son abordables mediante modelos de regresión tradicionales (Huettel et al. 2009; Mahmoudi et al. 2012; Lewis-Peacock, Norman 2013). Considere el siguiente ejemplo en el que el patrón de activación (ej., la señales de actividad en distintas áreas del cerebro) bajo dos condiciones o macroestados experimentales, llamados A y B , ha sido construido según la siguiente relación (reducida a dos voxels por simplicidad):

$$(1) \quad \begin{cases} A : voxel_1^2 + voxel_2^2 > r^2 \\ B : voxel_1^2 + voxel_2^2 < r^2 \end{cases}$$

Esto no es más que un círculo en el espacio de los voxels (espacio de estados o espacio fase, más generalmente). Si restringimos los valores a un intervalo finito y constante centrado en $(0, 0)$, llamado c , entonces podemos obligar a los macroestados A y B a poseer el mismo número de puntos; si y solo si r en la expresión (1) cumple que:

$$(2) \quad \frac{c^2}{2} = \pi r^2$$

Luego, se muestrean algunos puntos aleatoriamente de manera uniforme (40 en las figuras 3.1 y 3.2). Al graficarlos en el plano, se observa que nuestras mediciones contienen suficiente información para distinguir ambos estados cerebrales, incluso por simple inspección visual:

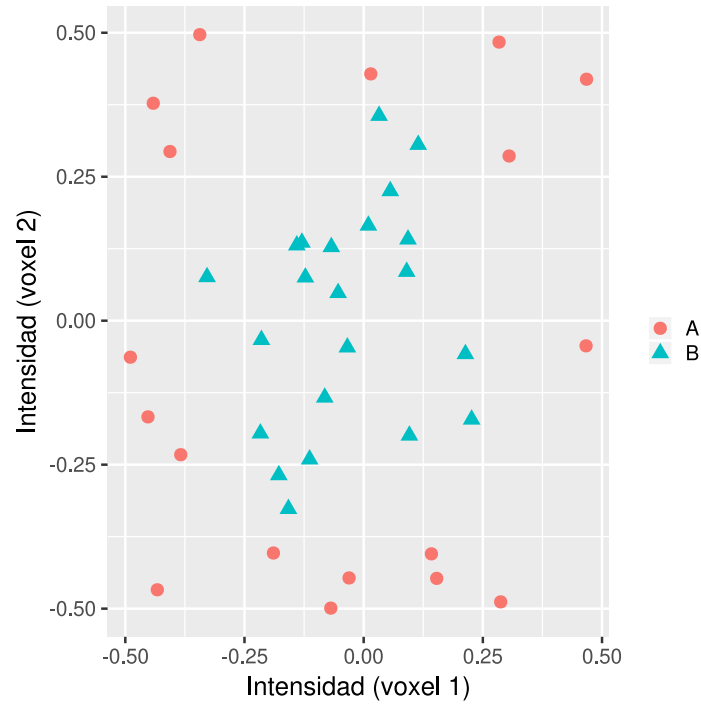


Figura 3.1: Patrón bidimensional no lineal de actividad de dos voxeles, según se muestreó aleatoriamente con 18 ensayos bajo la condición experimental *A* más 22 bajo la *B*. Las mediciones de la señales que corresponden a la condición *A* pueden distinguirse por su excentricidad con respecto de la condición *B*.

Sin embargo un análisis univariado (como regresión mediante GLM) está destinado a fracasar (ver figura 3.2). Las distribuciones unidimensionales de muestra se encuentran aproximadamente centradas alrededor del mismo valor de intensidad, así que no se observa ningún efecto intercondición. Por otro lado, las mediciones están muy dispersas y lucen ruidosas. Tal enormidad en el rango de valores para cada estado tampoco puede ser explicado por artefactos de movimiento ni por un proceso determinista no estacionario que empuje sistemáticamente los valores; ya que el proceso usado para generar los estados fue explícitamente ninguna de esas cosas.

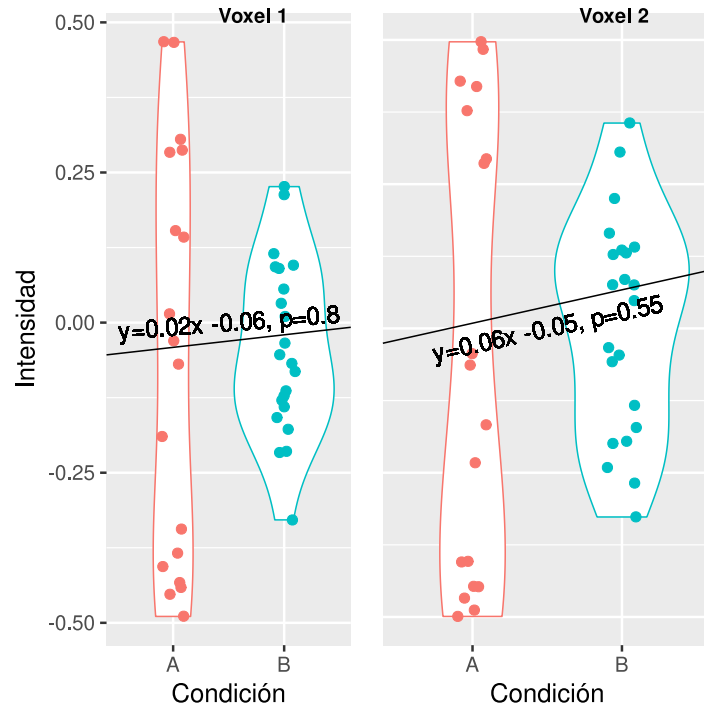


Figura 3.2: Modelos de regresión lineal simplificados y valores p de las pruebas t para las proyecciones de los datos de la figura 3.1 sobre cada eje. El efecto de las condiciones experimentales sobre voxels individuales es negligible (atestiguado por las pequeñas pendientes de los modelos), y probablemente son resultado del azar (valores p), como se esperaba. También por construcción, una prueba de Shapiro-Wilk revela que las distribuciones de datos A no pasan el criterio de normalidad, haciendo más cuestionable la aplicación de este tipo de análisis.

Incluso un clasificador univariado no lineal cometería una gran cantidad de errores en la intersección entre ambas distribuciones.¹ Un patrón discriminante univariado podría observarse a partir de un análisis dinámico del espectro de frecuencias bajo condiciones dinámicas especiales,² pero el enfoque más adecuado por mucho es considerar ambos voxels simultáneamente; como en la figura 3.1, ya que así es como se generaron los datos en primer lugar.

Paradójicamente, los estudios disponibles con MVPA han limitado sistemáticamente las bondades matemáticas de estos métodos; cuando evitan analizar datos de gran dimensionalidad espacial y cuando, pese a haber realizado un

¹Imagine que todos los puntos de un voxel son colocados sobre la recta real, y permitimos dibujar una curva continua sobre todo el plano, con tal de separar las dos condiciones.

²Si las oscilaciones de la trayectoria durante un estado fueran más rápidas, por ejemplo. O más generalmente, en la medida en que su descomposición de Fourier favorezca frecuencias exclusivas.

análisis de cerebro completo, pierden la oportunidad de realizar pruebas estadísticas rigurosas para derivar mapas anatómicos de actividad cerebral relacionada a la tarea. Ambas características son parte integral del análisis clásico, que los proponentes de MVPA no escatiman calificar como complementario, si no es que obsoleto. La gran mayoría de los estudios con MVPA más bien han enfatizado su capacidad para hacer decodificación neuronal y predicción a partir de un subconjunto de voxels o regiones de interés, seleccionadas *a priori* o heurísticamente, así se desperdicie la oportunidad de encontrar un mejor patrón haciendo una búsqueda guiada por los datos, más extensiva y distribuida (como son los correlatos cerebrales conocidos de las emociones).

Otras causas para estas flaquezas en la literatura de MVPA son más bien limitantes computacionales y metodológicas: la adición de variables puede ser detrimental para los resultados cuando sólo algunas de ellas contribuyen información y no ruido, además de complicar el cómputo. Por lo tanto puede ser legítimo a los objetivos de un estudio buscar reducir la dimensionalidad (también conocido como “selección de características”). Tanto la reducción de dimensiones como la selección de algoritmo de aprendizaje de patrones puede que impactare también en la factibilidad de producir mapas anatómicos de actividad, ya que muchos de ellos descartan la correspondencia entre sus propias variables y la topología de los voxels originales. Con todo y eso, éstas son áreas de gran oportunidad para los estudios de discriminación de emociones. De hecho, el algoritmo más recurrido (SVM) *sí tiene interpretación espacial* subutilizada. A continuación se revisarán los mejores acercamientos hasta ahora, seleccionando de la lista mencionada en el capítulo de *Antecedentes*.

En materia de percepción emocional, Kotz et al. (2012) hicieron pruebas estadísticas a nivel de grupo, produciendo mapas anatómicos de actividad relevante a la discriminación de expresión emocional en estímulos sonoros. Sin embargo, los datos de entrada al análisis grupal provienen de una búsqueda tipo “searchlight”, que se restringe a analizar voxels contiguos dentro de una esfera de pocos milímetros de radio. Skerry, Saxe (2015) también realizaron un análisis global mediante searchlight y análisis estadístico a nivel grupal, como justificación de sus ROIs y evidencia complementaria. Hernández-Pérez

et al. (2018) limitaron la elaboración de mapas anatómicos a un espacio de búsqueda con el 1 % de voxels más sobresalientes según sus puntajes F en una prueba ANOVA, y no se intentó un análisis grupal, probablemente porque el poco traslape espacial tras los ANOVAs por sujeto lo volvería fútil. Los puntajes F son una razón entre la varianza media intergrupala (Mean Square Among groups) y la varianza media intragrupal (Mean Square Within groups), como define la ecuación (3). Por lo tanto asume que los voxels tienen comportamientos gaussianos, y que los más relevantes para lograr la decodificación son los más separados del resto. Esto merma la decisión de usar MVPA, que suele ofrecer mayor generalidad.

(3)

$$F = \frac{MSA}{MSW} = \frac{\frac{1}{G-1} \sum_{g=1}^G n_g (\bar{x}_g - \bar{x}_{..})^2}{\frac{1}{N-G} \sum_{g=1}^G \sum_{i=1}^{n_g} (x_{gi} - \bar{x}_g)^2}$$

Donde G es el número de grupos, N el número neto de datos y n_g el número de datos para el grupo g .

En lo que respecta a experimentación de emociones en primera persona, Saarimäki et al. (2015) también filtraron sus voxels mediante puntajes F , luego derivaron mapas anatómicos para cada categoría emocional a partir de la actividad y los pesos de la red neuronal artificial entrenada para hacer la clasificación. Esta red únicamente contenía capa de entrada (con una neurona por voxel) y capa de salida (una neurona por categoría), por lo que su entrenamiento es equivalente a optimizar las combinaciones lineales de señales que “activan” una categoría; no muy distinto de lo que logra el algoritmo SVM lineal. Extrañamente, se reportó el producto del peso de la sinapsis artificial y el valor del voxel; asumiendo sin garantías que un voxel “silencioso” durante determinado estado emocional no contribuye a manufacturar dicho estado, aunque la red neuronal le hubiese dado mucho peso. Quizás el reporte más completo hasta ahora sea el de Kragel, LaBar (2015): un estudio multimodal de inducción de emociones que además comparó el modelo de emociones básicas vs un modelo dimensional. Utilizó un algoritmo de análisis discriminante por mínimos cuadrados parciales (PLS-DA) entrenado con los datos de todos los sujetos, de cuyos parámetros se derivaron (y valida-

ron estadísticamente mediante técnicas de remuestreo) patrones anatómicos para cada emoción.

Ademas, muchos de estos estudios podrían haber sido menos laxos en sus pruebas estadísticas a nivel grupal; por ejemplo, a causa del método de corrección de múltiples comparaciones, umbrales de significancia y uso de distribuciones paramétricas a priori que podrían no reflejar los verdaderos niveles de clasificación por azar. Así mismo, aunque algún tipo u otro de validación cruzada contra ejemplos fuera de la muestra de entrenamiento es estrictamente observada en la literatura; aquellos que recurrieron al esquema “leave-one-out”, o con “folds” que no corresponden a las adquisiciones de datos, no reportaron haber mitigado la autocorrelación de las señales. Es decir, si datos suficientemente temporalmente cercanos a los de entrenamiento terminaran en la partición de validación, podría ser que el buen desempeño del clasificador se debiera a similitudes en las tendencias de baja frecuencia y no a un aprendizaje de las características de interés. Esto es cierto en la medida que el funcional de correlación cruzada resulte alto.³⁴

³Se puede explicar mediante la siguiente analogía: imagine que toma consecutivamente dos fotografías de un amigo, con el afán de usar la primera para pedirle a un tercero que aprenda a reconocerlo, y la segunda para poner su habilidad a prueba. Puede que acierte debido a que en la segunda foto llevaba la misma ropa, porque el fondo y los colores son similares; y no debido a que aprendió a reconocer a la persona.

⁴Aunque no son usados explícitamente para este propósito, el uso de filtros pasa-altas y regresión para “detrending” podrían estar mitigando el problema.

Capítulo 4

Hipótesis

4.1. Preguntas de investigación

- ¿Es posible decodificar la emoción percibida en una expresión facial a partir de actividad fisiológica cerebral; en especial, considerando simultáneamente toda la información disponible?
- ¿Es posible obtener mapas estadístico-anatómicos de actividad cerebral relacionada a tarea mediante métodos multivariantes, considerando simultáneamente toda la información disponible?

4.2. Hipótesis

- Hipótesis nula: el poder predictivo de modelos de clasificación multivariada entrenados con neuroimágenes funcionales bajo percepción de dichos estímulos es *igual* al nivel de predicción por azar.
- Hipótesis de trabajo: el poder predictivo de modelos de clasificación multivariada entrenados con neuroimágenes funcionales bajo percepción de dichos estímulos es *mayor* al nivel de predicción por azar.

4.3. Predicciones

Si existiera evidencia (en las señales BOLD, por ejemplo) para falsar la hipótesis nula, entonces las estructuras anatómicas involucradas y su actividad hemodinámica podrían ser obtenidas a partir de los modelos de clasificación exitosos; ya sea a nivel individual o grupal.

Capítulo 5

Objetivos

5.1. General

- Estudiar las bases biológicas de la percepción emocional a escala de redes funcionales enteras de cerebros humanos.

5.2. Específicos

- Predecir la presentación de algunas emociones básicas a partir de datos de resonancia magnética funcional.
- Derivar imágenes de los correlatos neuroanatómicos y funcionales de esta facultad cognitiva, contribuir al campo de neurociencia afectiva con dicha evidencia.
- Construir y documentar un *pipeline* válido, general, reproducible y reutilizable para el análisis de datos de fMRI orientado a este tipo de tareas mediante clasificación multivariante de punta.
- Contrastar los resultados del método multivariante contra el análisis clásico por GLM.

Capítulo 6

Muestra, Materiales y Métodos

Se analizaron dos bases de datos distintas en búsqueda de evidencia de una representación neural de la emoción de los estímulos. Mientras que la primera proviene de un experimento psicológico que no involucró al autor del presente proyecto, el segundo experimento fue diseñado desde cero con el propósito expreso de mitigar las deficiencias del primero, en tanto que idóneo para responder a la pregunta de investigación. En lo subsecuente serán referidos como **Experimento 1** y **Experimento 2**, respectivamente. Es importante redundar en que ambos se destinaron a responder exactamente la misma cuestión, y por lo tanto serán descritos simultáneamente, resaltando sus diferencias.

6.1. Participantes

El Experimento 1 reunió 34 participantes neurotípicos; sin reporte de padecer o haber sido diagnosticados con alguna condición neurológica o psiquiátrica.

El Experimento 2 reunió 8 participantes, quienes también negaron padecer o en algún momento haber sido diagnosticados con alguna condición neurológica o psiquiátrica. Todos los sujetos fueron varones entre los 23 y 30 años de edad al momento del estudio. 7 de ellos reportaron tener fenotipo de lateralidad diestra, mientras que el restante reportó ser zurdo. Uno de los participantes diestros fue excluido del análisis debido a una falla en la adquisición de imágenes, ya que habría conducido a un desbalance en el número de muestras disponibles para entrenar los modelos de clasificación automática.

6.2. Estímulos y diseño experimental

Dado un rostro con mirada lateralizada, el sistema visual humano es inducido a redirigir la atención en la dirección sugerida; presuntamente porque información de relevancia ambiental y social podría ser descubierta en donde otros miran. Este cambio atencional abrupto se encuentra dentro del rango de intervalos rápidos de un movimiento sacádico. El fenómeno es conocido como “gaze cuing”, y ha sido usado para experimentar con la atención visual y su interacción con la percepción (Friesen, Kingstone 1998).

El Experimento 1 emplea una variación a la tarea de “gaze cuing” con un diseño por eventos. No obstante, para efectos de este trabajo el componente emocional de los estímulos puede considerarse como presentado en pequeños bloques de 2.7 segundos y 3 segundos, ya que es invariante dentro de cada ensayo. Cada sesión de registro comprende 126 ensayos estímulo-respuesta, donde ensayos con rostros que desvían la mirada son intercalados con otros de control (mirada directa) para evitar habituación en los sujetos y contaminación en el análisis de clasificación debido a correlación temporal. Esto se explica más a detalle en la figura 6.1: un diagrama de sistema de estados discretos de tipo markoviano, en el que se representa un par de bloques $\langle \textit{gaze cuing}, \textit{control} \rangle$. Varias iteraciones en el diagrama estarían ocurriendo durante el registro de imágenes.

Los verdaderos estímulos provienen de la base de datos fotográfica NimStim (Tottenham et al. 2009). En la ilustración han sido caricaturizados por brevedad, ya que se utilizó más de un rostro para cada emoción (4 hombres y 4 mujeres con variantes para todas las expresiones requeridas, étnicamente europeoides, presentados de forma balanceada. Los identificadores de los modelos dentro de la base de datos son: 3, 5, 7, 8, 24, 25, 32, 36).

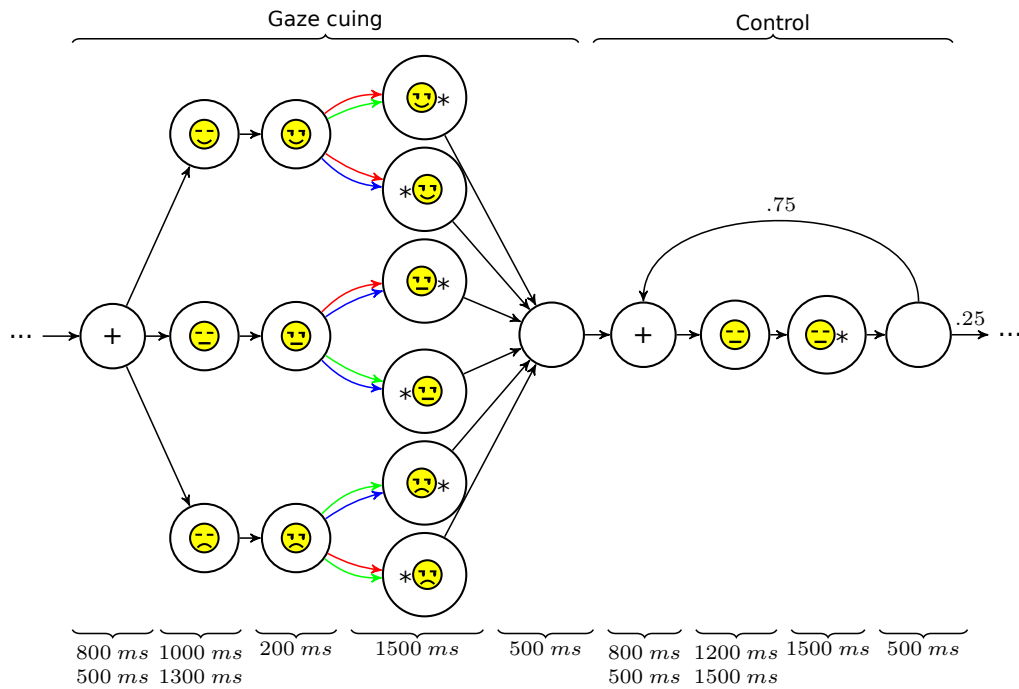


Figura 6.1: Paradigma de *gaze cuing*. Los nodos representan estímulos visuales, su orden de presentación es indicado usando flechas. A menos que se indique lo contrario, la probabilidad de transición desde un nodo se distribuye equitativamente entre todas las flechas salientes (los ensayos se presentaron pseudoaleatoriamente). Se le pidió a los participantes que buscaran un objetivo (asterisco) y que presionaran el botón izquierdo o derecho, según la posición del objetivo relativo a la cara. El espectro de expresiones probadas (neutral, feliz, triste; seguidas ya fuera por un objetivo congruente o incongruente con la mirada) fue dividido y contrabalanceado en tres secuencias, para mantener los ensayos suficientemente cortos. Las flechas coloridas indican combinaciones específicas a una de las tres secuencias.

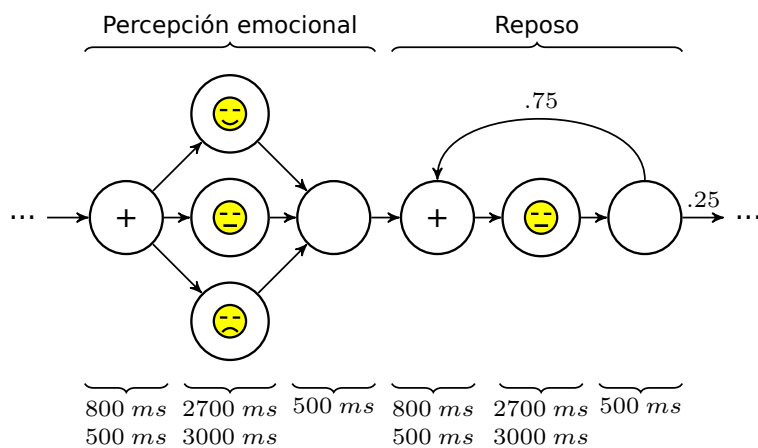


Figura 6.2: Para propósitos del etiquetado de datos de entrenamiento y prueba de clasificación mediante aprendizaje de máquinas, los estados de estimulación de la figura 6.1 pueden ser condensados en los metaestados emocionales aquí mostrados.

Para el Experimento 2 se utilizaron 5 secuencias contrabalanceadas en un diseño por bloques de 30 segundos. Las secuencias son idénticas salvo por el orden pseudoaleatorio de presentación de los bloques. A las 3 clases de bloques del Experimento 1 (neutral, felicidad, tristeza) se añadieron otras 3: de enojo, caras neutras revueltas como rompecabezas (“scrambled”) que conservan las mismas propiedades estadísticas de la imagen aunque ya no pueden ser reconocidas como caras, y bloques de estimulación visual mínima (“blank”). A su vez, cada bloque consta de la presentación aleatoria de 10 imágenes pertenecientes a esa categoría, durante aproximadamente 3 segundos cada una y sin posibilidad de repetirse dentro del mismo bloque. Cada bloque aparecía dos veces por secuencia, con lo que se obtiene un total de 12 bloques por secuencia ($360\text{ s} = 6\text{ min}$). Al término de los 12, toda estimulación y petición de respuesta era interrumpida, debiendo el sujeto esperar 10 segundos antes de concluir el experimento (para dar tiempo suficiente al escáner de capturar la respuesta hemodinámica elicitada por los últimos estímulos). Todas las secuencias del Experimento 2 se realizaron en una sola sesión por sujeto, de aproximadamente 30 minutos, 50 segundos ($370\text{ s} \times 5$), justo después de haber obtenido las imágenes anatómicas por resonancia magnética.

En lo que concierne a los estímulos que conforman cada bloque, éstos fueron seleccionados de la clásica base de datos de Paul Ekman, “Pictures of Facial Affect” (Ekman 1976). En el caso de los bloques de mínima estimulación visual (“blank”) simplemente se hacía “saltar” aleatoriamente una cruz de fijación tenue de un cuadrante a otro dentro del espacio visual. La cruz no ocupaba más del 10 % del espacio total en cualquiera de sus direcciones. Véase la figura 6.3 para una explicación gráfica del Experimento 2.

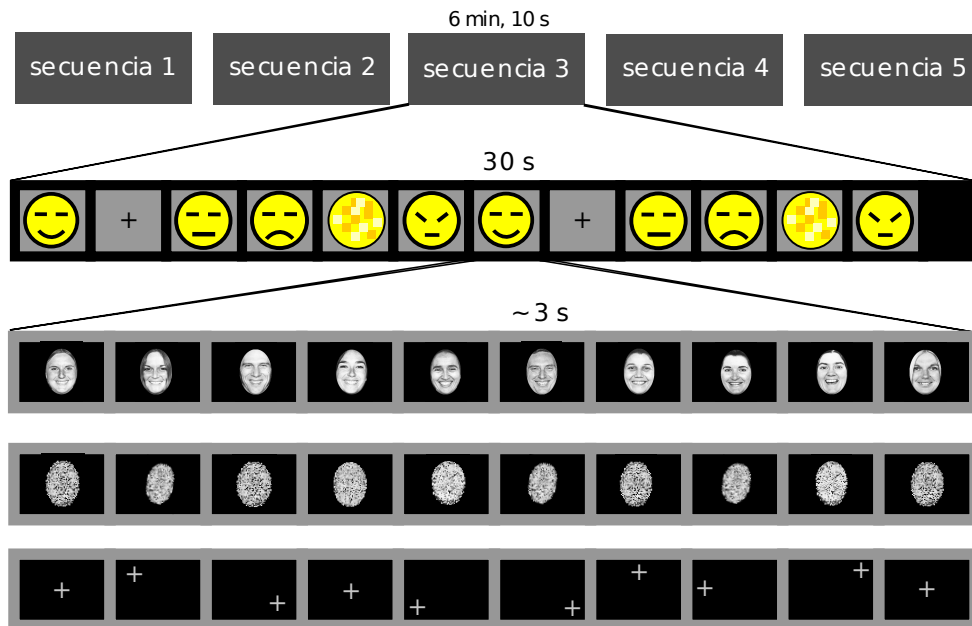


Figura 6.3: Paradigma por bloques del Experimento 2. El eje horizontal representa el paso del tiempo. Los rectángulos corresponden a unidades de estimulación (secuencias, bloques y estímulos individuales).

Se tomaron mediciones conductuales a través de la tarea como indicadores de la prestación de atención a los estímulos. Al inicio de cada secuencia los participantes recibieron la instrucción explícita de indicar si cada rostro presentado pertenecía a un hombre o a una mujer. Esto ocurría mediante botones para cada dedo índice, con controles a cada mano que les permitían mantener los brazos lo más fijos posibles. En cambio, cuando la imagen claramente no correspondía a un hombre o una mujer (o persona siquiera: bloques “scrambled” y “blank”), la instrucción simplemente era reportar cada cambio en la imagen usando los índices alternadamente. De esta manera se garantizaba que la actividad motora fuera la misma durante todos los bloques. Aunque jamás se les pidió catalogar o pensar en la expresión o emoción de los rostros, no es posible descartar la posibilidad de que este proceso lingüístico-conceptual apareciera espontáneamente en su tren de pensamiento; lo cual añadiría una variable de confusión al análisis de clasificación de patrones.

6.3. Adquisición de imágenes

Tanto las imágenes del primer como del segundo experimento fueron obtenidas con un resonador General Electric Discovery MR750 de la Unidad de Resonancia Magnética del Instituto de Neurobiología; cuyo campo magnético principal alcanza los $3,0\ T$ en su punto de magnitud máxima. Los ecos inducidos fueron registrados usando una antena de 32 canales montada a cabeza.

Los escaneos pesados a contraste $T2^*$ para imágenes funcionales durante el Experimento 1 se componen de 8840 cuadros cada uno. La tabla 6.1 describe sus parámetros de secuencia.

Cuadro 6.1: Parametros de secuencia utilizados durante las adquisiciones de IRMf.

Parámetro de secuencia	Valor
Tipo de adquisición	2D
Orientación de rebanadas	Transversal
Rebanadas	34
Tamaño de matriz	64×64
Tamaño de voxel	$(4\ mm)^3$
Flip angle	$\pi/2$
TR	$2000\ ms$
TE	$40\ ms$

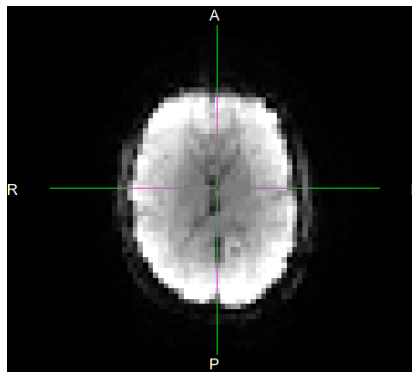


Figura 6.4: Muestra de imagen eco-planar cruda dentro del conjunto de datos. Este corte axial es una de las 34 rebanadas adquiridas dentro de cada tiempo de repetición (TR).

6.4. Métodos de análisis

Capítulo 7

Resultados

Capítulo 8

Discusión

Capítulo 9

Conclusiones

Apéndice 1: código fuente

Listing 9.1: Modelos predictivos de un solo voxel vs multivoxel. El siguiente programa en R fue utilizado para generar las figuras 3.2 y 3.1. Además se incluye una breve demostración del clasificador SVM.

```
1 library(ggplot2)
2 theme_set(theme_gray(base_size = 18))
3 library(cowplot)
4 library(e1071) # SVM classifier
5
6 set.seed(111)
7 N <- 40
8 CONDITIONS <- c("A", "B")
9
10 ## uniformly sample N points within (0,1)x(0,1)
11 data <- data.frame(replicate(2, runif(N, min = -.5, max = .5)))
12 colnames(data) <- c("voxel_1", "voxel_2")
13
14 ## divide space into 2 condition regions according to some boundary relation
15 label <- function(p) {
16   ## showcase nonlinear capabilities: circle of area .5 centered at (0, 0)
17   diameter <- .5 / pi
18   if ((p["voxel_1"])**2 > diameter - (p["voxel_2"])**2) {
19     CONDITIONS[1]
20   } else {
21     CONDITIONS[2]
22   }
23 }
24 data <- cbind(data, cond = apply(data, 1, label)) # label each point
25 ## reorder according to label, for visual convenience
26 data <- data[with(data, order(cond)), ]
27
28 ## plot single-voxel models
29 lm_plot <- function(data, xname, yname, remove_ytext = FALSE) {
30   model <- summary(lm(paste0(yname, "~", xname),
31     data))
32   beta1 <- model$coefficients[paste0(xname, "B"), "Estimate"]
33   beta0 <- model$coefficients["(Intercept)", "Estimate"]
34   p <- model$coefficients[paste0(xname, "B"), "Pr(>|t|)"]
35   plot <- ggplot(data, aes(x = data[, xname],
36     y = data[, yname],
37     color = data[, xname])) +
38     labs(x = "Condición", y = "Intensidad") +
39     geom_violin() +
40     geom_jitter(width = .1, size = 3) +
41     geom_abline(slope = beta1,
42       intercept = beta0) +
43     geom_text(color = "black",
44       angle = atan(beta1) * (180 / pi) * 3.75,
45       aes(x = 1.5,
46         y = 0,
47         label = paste0("y=",
48           round(beta1, 2),
49           "x ")),
```



```

50         round(beta0, 2),
51         ", p=",
52         round(p, 2))),
53     size = 6) +
54     theme(legend.position="none")
55   if (remove_ytext) {
56     plot <- plot + theme(axis.title.y = element_blank(),
57       axis.text.y = element_blank(),
58       axis.ticks.y = element_blank())
59   }
60   shapiro.test(data[data$cond == CONDITIONS[1], as.character(yname)])
61   shapiro.test(data[data$cond == CONDITIONS[2], as.character(yname)])
62   plot
63 }
64
65 svg("./lm-vs-multivariate-1.svg")
66 cowplot::plot_grid(lm_plot(data, xname = "cond", yname = "voxel_1"), # no effect, p
  = .795
67   lm_plot(data, xname = "cond", yname = "voxel_2", TRUE), # no effect,
  p = .546
68   labels = c("Voxel 1", "Voxel 2"),
69   label_x = .5)
70 dev.off()
71
72 ## summary(lm(formula = cond ~ voxel_1:voxel_2, data = data)) # no interaction
73
74 svg("./lm-vs-multivariate-2.svg")
75 ggplot(data, aes(x = voxel_1, y = voxel_2, color = cond, shape = cond)) +
76   geom_point(size = 4) +
77   labs(x = "Intensidad (voxel 1)",
78     y = "Intensidad (voxel 2)",
79     color = "Condition") +
80   scale_color_discrete("") +
81   scale_shape_manual("", values = c(16, 17))
82 dev.off()
83
84 indices <- sample(1:nrow(data), N/2)
85 model3 <- svm(cond ~ voxel_2 + voxel_1,
86   data = data[indices, ],
87   kernel = "radial",
88   cost = 1000,
89   gamma = 0.01)
90 predict(model3, data[-indices, ])
91 svg("./lm-vs-multivariate-3.svg")
92 plot(model3, data)
93 dev.off()

```

Referencias

- American Psychiatric Association, 2013. *Diagnostic and statistical manual of mental disorders (DSM-5)*, American Psychiatric Pub.
- Barrett, L.F., 2006. Are emotions natural kinds? *Perspectives on psychological science*, 1(1), pp.28–58.
- Barrett, L.F., Wager, T.D., 2006. The structure of emotion: Evidence from neuroimaging studies. *Current Directions in Psychological Science*, 15(2), pp.79–83.
- Baucom, L.B. et al., 2012. Decoding the neural representation of affective states. *Neuroimage*, 59(1), pp.718–727.
- Bechara, A., Damasio, A.R., Damasio, H., 2001. Insensitivity to future consequences following damage to human prefrontal. *The Science of Mental Health: Personality and personality disorder*, 50, p.287.
- Bechara, A., Damasio, H., Damasio, A.R., 2000. Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex*, 10(3), pp.295–307.
- Berrios, G.E., 1993. European views on personality disorders: A conceptual history. *Comprehensive Psychiatry*, 34(1), pp.14–30.
- Calder, A.J., Lawrence, A.D., Young, A.W., 2001. Neuropsychology of fear and loathing. *Nature Reviews Neuroscience*, 2(5), p.352.
- Celeghin, A. et al., 2017. Basic emotions in human neuroscience: Neuroimaging and beyond. *Frontiers in Psychology*, 8, p.1432.
- Chalmers, D.J., 1995. Facing up to the problem of consciousness. *Journal of consciousness studies*, 2(3), pp.200–219.
- Chang, L.J. et al., 2015. A sensitive and specific neural signature for picture-induced negative affect. *PLoS biology*, 13(6), p.e1002180.
- Chikazoe, J. et al., 2014. Population coding of affect across stimuli, modalities and individuals. *Nature neuroscience*, 17(8), p.1114.
- Damasio, A., Carvalho, G.B., 2013. The nature of feelings: Evolutionary and neurobiological origins. *Nature reviews neuroscience*, 14(2), p.143.
- Damasio, A.R., 1998. Emotion in the perspective of an integrated nervous system. *Brain research reviews*, 26(2-3), pp.83–86.

- Damasio, A.R., 1999. *The feeling of what happens: Body and emotion in the making of consciousness*, Houghton Mifflin Harcourt.
- Damasio, A.R., 1996. The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 351(1346), pp.1413–1420.
- Donath, J., 2007. Signals, cues and meaning. *Unpublished Manuscript. Massachusetts Institute of Technology, Cambridge, MA*. Available at: <http://smg.media.mit.edu/classes/IdentitySignals06/SignalingDraft.pdf>.
- Ekman, P., 2003. Darwin, deception, and facial expression. *Annals of the New York Academy of Sciences*, 1000(1), pp.205–221.
- Ekman, P., 1976. Pictures of facial affect. *Consulting Psychologists Press*.
- Ekman, P.E., Davidson, R.J., 1994. *The nature of emotion: Fundamental questions.*, Oxford University Press.
- Ellis, H.D. et al., 1997. Reduced autonomic responses to faces in capgras delusion. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 264(1384), pp.1085–1092.
- Ethofer, T. et al., 2009. Decoding of emotional information in voice-sensitive cortices. *Current Biology*, 19(12), pp.1028–1033.
- Fessler, D.M.T., Gervais, M., 2010. From whence the captains of our lives: Ultimate and phylogenetic perspectives on emotions in humans and other primates. In P. M. Kappeler, J. Silk, eds. *Mind the gap: Tracing the origins of human universals*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 261–280. Available at: https://doi.org/10.1007/978-3-642-02725-3_12.
- Fridlund, A.J., 2014. *Human facial expression: An evolutionary view*, Academic Press.
- Friesen, C.K., Kingstone, A., 1998. The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic bulletin & review*, 5(3), pp.490–495.
- Friston, K., 2010. The free-energy principle: A unified brain theory? *Nature reviews neuroscience*, 11(2), p.127.
- Friston, K.J. et al., 1994. Statistical parametric maps in functional imaging: A general linear approach. *Human brain mapping*, 2(4), pp.189–210.
- Gazzaniga, M., Ivry, R.B., 2013. *Cognitive neuroscience: The biology of the mind: Fourth international student edition*, WW Norton.
- Guillory, S.A., Bujarski, K.A., 2014. Exploring emotions using invasive methods: Review

- of 60 years of human intracranial electrophysiology. *Social cognitive and affective neuroscience*, 9(12), pp.1880–1889.
- Haist, F., Anzures, G., 2017. Functional development of the brain’s face-processing system. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(1-2), p.e1423.
- Hamann, S., 2012. Mapping discrete and dimensional emotions onto the brain: Controversies and consensus. *Trends in cognitive sciences*, 16(9), pp.458–466.
- Harrison, N.A. et al., 2010. The embodiment of emotional feelings in the brain. *Journal of Neuroscience*, 30(38), pp.12878–12884.
- Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. *Trends in cognitive sciences*, 4(6), pp.223–233.
- Haynes, J.-D., Rees, G., 2005. Predicting the stream of consciousness from activity in human visual cortex. *Current Biology*, 15(14), pp.1301–1307.
- Hernández-Pérez, R., Concha, L., Cuaya, L.V., 2018. Decoding human emotional faces in the dog’s brain. *bioRxiv*, p.134080.
- Huettel, S.A. et al., 2009. *Functional magnetic resonance imaging* 2nd ed., Sinauer Associates Sunderland, MA.
- Hurlemann, R. et al., 2007. Amygdala control of emotion-induced forgetting and remembering: Evidence from Urbach-Wiethe disease. *Neuropsychologia*, 45(5), pp.877–884.
- Isen, A.M., Patrick, R., 1983. The effect of positive feelings on risk taking: When the chips are down. *Organizational behavior and human performance*, 31(2), pp.194–202.
- Jimura, K., Poldrack, R.A., 2012. Analyses of regional-average activation and multivoxel pattern information tell complementary stories. *Neuropsychologia*, 50(4), pp.544–552.
- Kassam, K.S. et al., 2013. Identifying emotions on the basis of neural activation. *PloS one*, 8(6), p.e66032.
- Kober, H. et al., 2008. Functional grouping and cortical–subcortical interactions in emotion: A meta-analysis of neuroimaging studies. *Neuroimage*, 42(2), pp.998–1031.
- Kotz, S.A. et al., 2012. Predicting vocal emotion expressions from the human brain. *Human Brain Mapping*, 34(8), pp.1971–1981.
- Kragel, P.A., LaBar, K.S., 2014. Advancing emotion theory with multivariate pattern classification. *Emotion Review*, 6(2), pp.160–174.
- Kragel, P.A., LaBar, K.S., 2016. Decoding the nature of emotion in the brain. *Trends in cognitive sciences*, 20(6), pp.444–455.

- Kragel, P.A., LaBar, K.S., 2015. Multivariate neural biomarkers of emotional states are categorically distinct. *Social cognitive and affective neuroscience*, 10(11), pp.1437–1448.
- Laidre, M.E., Johnstone, R.A., 2013. Animal signals. *Current Biology*, 23(18), pp.R829–R833.
- LeDoux, J.E., Damasio, A.R., 2013. In E. R. Kandel et al., eds. *Principles of neural science*. McGraw-hill New York.
- Lewis, T., 2013. Oldest fossils reveal when apes & monkeys first diverged. Available at: <https://www.livescience.com/32029-oldest-monkey-fossil-found.html>.
- Lewis-Peacock, J.A., Norman, K.A., 2013. Multi-voxel pattern analysis of fMRI data. *The cognitive neurosciences*, pp.911–920.
- Lindquist, K.A., Barrett, L.F., 2012. A functional architecture of the human brain: Emerging insights from the science of emotion. *Trends in cognitive sciences*, 16(11), pp.533–540.
- Lindquist, K.A. et al., 2012. The brain basis of emotion: A meta-analytic review. *Behavioral and brain sciences*, 35(3), pp.121–143.
- Logothetis, N.K., 2008. What we can do and what we cannot do with fMRI. *Nature*, 453(7197), p.869.
- Mahmoudi, A. et al., 2012. Multivoxel pattern analysis for fmri data: A review. *Computational and mathematical methods in medicine*, 2012.
- Marlowe, W.B., Mancall, E.L., Thomas, J.J., 1975. Complete Klüver-Bucy syndrome in man. *Cortex*, 11(1), pp.53–59.
- Michel, M. et al., 2019. Opportunities and challenges for a maturing science of consciousness. *Nature human behaviour*, 3(2), p.104.
- Millon, T., Davis, R.O., 1996. *Disorders of personality: DSM-iv and beyond*, John Wiley & Sons.
- Mitchell, R.L., Phillips, L.H., 2015. The overlapping relationship between emotion perception and theory of mind. *Neuropsychologia*, 70, pp.1–10.
- Murphy, F.C., Nimmo-Smith, I., Lawrence, A.D., 2003. Functional neuroanatomy of emotions: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, 3(3), pp.207–233.
- Newen, A., De Bruin, L., Gallagher, S., 2018. *The oxford handbook of 4E cognition*, Oxford University Press.
- Norman, K.A. et al., 2006. Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in cognitive sciences*, 10(9), pp.424–430.

- Oizumi, M., Albantakis, L., Tononi, G., 2014. From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0. *PLoS computational biology*, 10(5), p.e1003588.
- Peelen, M.V., Atkinson, A.P., Vuilleumier, P., 2010. Supramodal representations of perceived emotions in the human brain. *Journal of Neuroscience*, 30(30), pp.10127–10134.
- Penny, W.D. et al., 2011. *Statistical parametric mapping: The analysis of functional brain images*, Elsevier.
- Pessoa, L., Padmala, S., 2007. Decoding near-threshold perception of fear from distributed single-trial brain activation. *Cerebral cortex*, 17(3), pp.691–701.
- Phan, K.L. et al., 2002. Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, 16(2), pp.331–348.
- Polyn, S.M. et al., 2005. Category-specific cortical activity precedes retrieval during memory search. *Science*, 310(5756), pp.1963–1966.
- Porter, S., Ten Brinke, L., 2008. Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions. *Psychological science*, 19(5), pp.508–514.
- Ramachandran, V., 1998. Consciousness and body image: Lessons from phantom limbs, capgras syndrome and pain asymbolia. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 353(1377), pp.1851–1859.
- Rissman, J., Wagner, A.D., 2012. Distributed representations in memory: Insights from functional brain imaging. *Annual review of psychology*, 63, pp.101–128.
- Rohrer, J.D. et al., 2011. Clinical and neuroanatomical signatures of tissue pathology in frontotemporal lobar degeneration. *Brain*, 134(9), pp.2565–2581.
- Rolls, E.T., Grabenhorst, F., Franco, L., 2009. Prediction of subjective affective state from brain activations. *Journal of Neurophysiology*, 101(3), pp.1294–1308.
- Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage*, 52(3), pp.1059–1069.
- Russell, J.A., Dols, J.M.F., 1997. *The psychology of facial expression*, Cambridge university press Cambridge.
- Saarimäki, H. et al., 2015. Discrete neural signatures of basic emotions. *Cerebral cortex*, 26(6), pp.2563–2573.
- Said, C.P. et al., 2010. Distributed representations of dynamic facial expressions in the superior temporal sulcus. *Journal of vision*, 10(5), pp.11–11.
- Sander, D., Grafman, J., Zalla, T., 2003. The human amygdala: An evolved system for

- relevance detection. *Reviews in the Neurosciences*, 14(4), pp.303–316.
- Schacter, D.L., Gilbert, D.T., Wegner, D.M., 2011. *Psychology (2nd edition)*, New York: Worth; Worth.
- Schirmer, A., Adolphs, R., 2017. Emotion perception from face, voice, and touch: Comparisons and convergence. *Trends in cognitive sciences*, 21(3), pp.216–228.
- Schwarz, N., 1990. Feelings as information: Informational and motivational functions of affective states. In T. E. Higgins, R. M. Sorrentino, eds. *Handbook of motivation and cognition: Foundations of social behavior*. Guilford Press, pp. 527–561.
- Schwarz, N., Clore, G.L., 2003. Mood as information: 20 years later. *Psychological inquiry*, 14(3-4), pp.296–303.
- Seth, A.K., 2013. Interoceptive inference, emotion, and the embodied self. *Trends in cognitive sciences*, 17(11), pp.565–573.
- Seth, A.K., 2010. The grand challenge of consciousness. *Frontiers in psychology*, 1, p.5.
- Shamay-Tsoory, S.G., Aharon-Peretz, J., Perry, D., 2009. Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, 132(3), pp.617–627.
- Shinkareva, S.V. et al., 2014. Representations of modality-specific affective processing for visual and auditory stimuli derived from functional magnetic resonance imaging data. *Human brain mapping*, 35(7), pp.3558–3568.
- Simon, D. et al., 2008. Recognition and discrimination of prototypical dynamic expressions of pain and emotions. *PAIN*, 135(1-2), pp.55–64.
- Simon, D. et al., 2006. Brain responses to dynamic facial expressions of pain. *Pain*, 126(1-3), pp.309–318.
- Sitaram, R. et al., 2011. Real-time support vector classification and feedback of multiple emotional brain states. *Neuroimage*, 56(2), pp.753–765.
- Skerry, A.E., Saxe, R., 2015. Neural representations of emotion are organized around abstract event features. *Current biology*, 25(15), pp.1945–1954.
- Slaby, J., 2014. Emotions and the extended mind. *Collective emotions*, pp.32–46.
- Soon, C.S. et al., 2008. Unconscious determinants of free decisions in the human brain. *Nature neuroscience*, 11(5), p.543.
- Tong, F., Pratte, M.S., 2012. Decoding patterns of human brain activity. *Annual review of psychology*, 63, pp.483–509.

- Tononi, G., Koch, C., 2015. Consciousness: Here, there and everywhere? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1668), p.20140167.
- Torralva, T. et al., 2009. A neuropsychological battery to detect specific executive and social cognitive impairments in early frontotemporal dementia. *Brain*, 132(5), pp.1299–1309.
- Tottenham, N. et al., 2009. The nimstim set of facial expressions: Judgments from untrained research participants. *Psychiatry research*, 168(3), pp.242–249.
- Tracy, J.L., Randles, D., Steckler, C.M., 2015. The nonverbal communication of emotions. *Current opinion in behavioral sciences*, 3, pp.25–30.
- Tsao, D.Y. et al., 2006. A cortical region consisting entirely of face-selective cells. *Science*, 311(5761), pp.670–674.
- Tsao, D.Y., Livingstone, M.S., 2008. Mechanisms of face perception. *Annu. Rev. Neurosci.*, 31, pp.411–437.
- Tsuchiya, N., Adolphs, R., 2007. Emotion and consciousness. *Trends in cognitive sciences*, 11(4), pp.158–167.
- Vytal, K., Hamann, S., 2010. Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis. *Journal of cognitive neuroscience*, 22(12), pp.2864–2885.
- Wager, T.D. et al., 2008. The neuroimaging of emotion. *Handbook of emotions*, 3, pp.249–271.
- Wegrzyn, M. et al., 2015. Investigating the brain basis of facial expression perception using multi-voxel pattern analysis. *Cortex*, 69, pp.131–140.
- World Health Organization, 2004. *International statistical classification of diseases and related health problems* 10th ed., World Health Organization.
- Zaki, J., Ochsner, K.N., 2012. The neuroscience of empathy: Progress, pitfalls and promise. *Nature neuroscience*, 15(5), p.675.