DOCTOR OF PHILOSOPHY

Investigating the neurocognitive bases of social cognition from a general semantics perspective

Diveica, Veronica

Award date:
2023

Awarding institution:
Bangor University

Investigating the neurocognitive bases of social cognition from a general semantics perspective

Veronica Diveică

Thesis submitted to the School of Human and Behavioural Sciences, Bangor University, in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Bangor, United Kingdom

December 2022

Declaration

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. All other sources are acknowledged by bibliographic references. This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless, as agreed by the University, for approved dual awards.

I confirm that I am submitting this work with the agreement of my Supervisor(s).

_____

Yr wyf drwy hyn yn datgan mai canlyniad fy ymchwil fy hun yw'r thesis hwn, ac eithrio lle nodir yn wahanol. Caiff ffynonellau eraill eu cydnabod gan droednodiadau yn rhoi cyfeiriadau eglur. Nid yw sylwedd y gwaith hwn wedi cael ei dderbyn o'r blaen ar gyfer unrhyw radd, ac nid yw'n cael ei gyflwyno ar yr un pryd mewn ymgeisiaeth am unrhyw radd oni bai ei fod, fel y cytunwyd gan y Brifysgol, am gymwysterau deuol cymeradwy.

Rwy'n cadarnhau fy mod yn cyflwyno'r gwaith hwn gyda chytundeb fy Ngoruchwyliwr (Goruchwylwyr)

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr Richard Binney. Richard's mentorship, guidance and words of encouragement have kept me on track and allowed me to develop the confidence and abilities needed to complete this PhD.

~

I would also like to thank all my collaborators. I am especially grateful to Prof Penny Pexman for her enthusiasm, mentorship, and willingness to impart her expert knowledge; to Dr Rebecca Jackson for helpful conversations and her insightful feedback; and to my second supervisor, Dr Kami Koldewyn, for her support and encouragements.

~

Additional thanks go to all the PhD students at Bangor University who have shared in this experience, especially to Eva, Ionela and Paul, for their support and all the stimulating conversations.

~

I want to thank my friends and cheerleaders, Andreea, Dana and Moneeb, for keeping my spirits high throughout this process.

~

I am eternally grateful for my family, whose unconditional love has been and will always be my primary driving force. I am deeply indebted to my parents for encouraging me to pursue this ambitious and life-changing educational journey. This thesis would not exist without their unceasing support and their belief in me.

~

I am also grateful to the ESRC Wales Doctoral Training Partnership for funding my PhD studentship.

Summary

Social interactions pervade every aspect of our lives and, thus, understanding their underlying cognitive mechanisms and brain basis is a major aim for cognitive neuroscience. Historically, the dominant approach within the social neuroscience sub-field has been to identify discrete systems that are tuned for processing information of a social nature. Recently, there has been growing interest in whether understanding the role of more general cognitive mechanisms can open up new perspectives and research questions in social neuroscientific research. This thesis adopts one such generalist approach and seeks to evaluate whether studying social cognition from a general semantics perspective can provide clues about its neurocognitive bases. In this view, social cognition is underpinned by (i) a system that represents conceptual knowledge and (ii) a control system that enables the task-appropriate use of this knowledge. Accordingly, the current work investigates the contribution of these two systems to social information processing.

Chapter 2 synthesises evidence in favour of the proposal that social semantic content, or *socialness*, makes unique contributions to conceptual representation, and highlights methodological issues, including the heterogeneity and inconsistency of the definitions used to quantify this construct. To address this limitation, Chapter 3 employs a novel and inclusive definition to collect the largest set of socialness ratings available to date, encompassing norms for over 8,000 individual concepts. Using these norms, the relationship between socialness and (i) other established semantic dimensions and (ii) behavioural indices of lexical-semantic processing are assessed. The results support the proposal that socialness is a distinct and behaviourally-relevant aspect of conceptual representation.

Chapter 4 reports a set of meta-analyses of functional neuroimaging data spanning 499 experiments and over 12,000 participants. It directly tests the prediction that social and semantic cognition rely on shared neural correlates, focusing on the semantic control network. The results show that a diverse range of social tasks, probing either mental state inference, trait inference, empathy or moral reasoning, reliably recruit brain regions that are sensitive to increased semantic control demands, and particularly the left inferior frontal gyrus. These findings are consistent with the proposal that social cognition is regulated, at least in part, by mechanisms for the controlled retrieval and selection of task-appropriate conceptual knowledge.

Chapter 5 explores the functional organisation of the left inferior frontal gyrus by means of a large-scale bimodal and data-driven analysis of its functional connectivity patterns. The

results are indicative of gradual variation in function across two principal spatial dimensions, and suggest that this region makes dissociable contributions to both social and non-social tasks, by virtue of differential affiliations with distinct large-scale functional networks. Our findings further suggest that the left inferior frontal gyrus acts as an interface between domain-general control and semantic systems, and this might make it ideally suited for implementing control processes on conceptual representations.

In summary, this body of work provides novel insights into the neurocognitive basis of social cognition, demonstrating the utility of studying social information processing through a general semantics lens.

Table of contents

List of figures

List of tables

This thesis has been prepared in alternative format meaning that Chapters 3-5 are written in the style of academic journal articles. Each of these Chapters is self-contained and includes a review of literature relevant to the work conducted, a description of the methods and results, and a discussion of the implications of the findings. The introductory Chapters 1-2 provide the broader theoretical context behind all the work in the thesis and identify the key research questions and objectives. Chapter 6 provides a general discussion that brings together and synthesizes the implications of this body of work as a whole.

CHAPTER 1

General Introduction

Perhaps the most ambitious quest the human mind has ever embarked on is understanding itself. To this end, the field of cognitive neuroscience investigates how the brain gives rise to thoughts and behaviours using experimental procedures from cognitive psychology in tandem with neuroimaging and neurostimulation techniques. In the past three decades, this approach has generated an explosion in neurobiological models of cognition, casting new light on sensory processing, as well as complex cognitive functions, like attention, memory and executive functioning (Gazzaniga, 2014; Poeppel et al., 2020). Brain function has traditionally been studied in isolation and by conceptualising the individual as a processor of information (Rescorla, 2015). However, a central aspect of life is that we do not exist in isolation; social interactions pervade every aspect of our existence. Consequently, the ability to comprehend and respond appropriately to the thoughts and actions of others is essential to thrive in society. This is evidenced by the profound detrimental effects that social isolation and social deficits caused by brain damage have on everyday functioning and quality of life (Holt-Lunstad et al., 2015; McDonald, 2013; Rosema et al., 2012). Therefore, understanding the neurocognitive mechanisms underpinning social interactions constitutes a major research challenge, which has led to the development of a somewhat distinct social neuroscientific field.

The dominant approach within social neuroscience has so far been to delineate systems that are tailored towards processing information of a social nature, the so-called 'domain-specific' approach (Ramsey & Ward, 2020). This has been borne out of an argument that the increasing complexity of social life has driven an expansion of the brain throughout primate evolution (Brothers, 1990; Dunbar, 1998). This 'social brain' hypothesis has led to a field-defining assumption that there are brain regions uniquely recruited by social cognitive processes. It is true that there is now mounting evidence to support the existence of cortical regions that respond preferentially to socially-relevant stimuli, such as faces (Kanwisher & Yovel, 2006), bodies (Peelen & Downing, 2007), biological motion (Grossman et al., 2000) and visual scenes depicting interactions between social agents (Isik et al., 2017). However, whether this purported specialization can be observed at the level of higher-order socio-cognitive functions, such as mental state inference and social regulation, is still debated (e.g., Binney & Ramsey, 2020; Ramsey & Ward, 2020; Spunt & Adolphs, 2017). As such, the

cognitive structure of social cognition is still underspecified (e.g., Alcalá-López et al., 2019; Happé et al., 2017) and a neurocognitive model that provides a unifying account of multiple social phenomena in the context of both health and disease is yet to be developed.

Recently, some researchers have argued that greater consideration needs to be given to the role of more generalizable neurocognitive systems in social cognition (Amodio, 2019; Binney & Ramsey, 2020; Heyes, 2014; Kilner, 2011; Michael & D'Ausilio, 2015; Ramsey, 2018; Ramsey & Ward, 2020; Spreng, 2013; Zaki, 2013). By definition, domain-general systems, such as those underpinning executive control, operate to some degree across all stimuli and tasks (Barrett, 2012), so their contribution should extend to contexts of a social nature[1]. Accordingly, this thesis investigates social cognition through a domain-general lens. In the following sections of this chapter, I aim to motivate this approach by synthesising relevant literature. First, I intend to demonstrate that researchers expect too much explanatory power from domain-specific systems alone by referencing findings from the mirror neuron and theory of mind literatures. Then, I outline the value of studying social cognition from a domain-general perspective. I argue that domain-general approaches have great potential to 1) lead to a more comprehensive understanding of social cognition and its relationship to generalizable mechanisms, 2) provide a unifying framework that can account for multiple forms of social cognition, and 3) act as a roadmap for future social neuroscientific research. To this end, I place social cognition within a semantic framework that treats social information processing as one example of how meaning is extracted from our everyday sensory experiences. Then, I provide an overview of an established model of semantic cognition, the controlled semantic cognition framework (Lambon Ralph et al., 2017), which is supported by a wealth of converging multimodal evidence that can be leveraged to generate new avenues for research into social cognition. In the last section of this chapter, I identify a set of outstanding questions about the role of conceptual knowledge and cognitive control in social cognition that will be addressed in the subsequent empirical chapters.

---

[1] There are various definitions of domain-specificity and domain-generality (for related discussions, see Barrett, 2012; Spunt & Adolphs, 2017). In this thesis, I generally use the term *domain-general* to refer to cognitive/brain mechanisms that operate on both social and non-social stimuli and tasks. In contrast, the term *domain-specific* is used to refer to cognitive/brain mechanisms specialized for processing social stimuli and tasks. However, domain-generality does not imply there is no form of functional specialization. For example, domain-general systems can be specialized for executive control and semantic operations. Relatedly domain-specificity is distinguishable from *modality-specificity*, which refers to specialization for processing inputs associated with a specific sensory modality (e.g., vision, audition, touch), regardless of their social or non-social nature.

1. Domain-specific systems alone have limited explanatory power

To date, an implicit assumption in many social frameworks is that the underlying neurocognitive systems differ between social and non-social contexts. As such, social neuroscientific studies tend to focus on the differences between social and non-social tasks and stimuli, while their similarities are relatively neglected. Consequently, the potential contribution of domain-general systems that operate across contexts is downplayed, and too much explanatory power is expected from the operation of domain-specific systems alone (Ramsey & Ward, 2020). The following sub-sections provide two examples of how the over-reliance on domain-specific explanations can manifest.

1.1. The mirror neuron mechanism

The mirror neuron mechanism is a prime example of how researchers' expectations for what domain-specific systems alone can explain are sometimes over-optimistic. 'Mirror neurons' refer to neurons that fire during both the execution and observation of an action. This type of neuron was first discovered using single-unit recordings in the premotor cortex of the macaque brain (di Pellegrino et al., 1992; Gallese et al., 1996) and later observed in humans (Mukamel et al., 2010). The mirror neuron network, defined using non-invasive functional neuroimaging as brain areas that increase their BOLD response during both action execution and observation, includes the premotor cortex, inferior frontal gyrus (IFG), and inferior parietal lobule (Figure 1A; Molenberghs et al., 2012). Although these regions display mirroring properties (Bonini et al., 2022; Chong et al., 2008; Kilner et al., 2009), this pattern of BOLD responses is not necessarily caused by mirror neuron activity, so these regions are sometimes called the 'action observation network'.

Given that mirror neurons provide a brain mechanism for linking 'my action' and 'your action', there was an expectation that their discovery would represent a "revolution" in the study of social behaviour (Iacoboni, 2008). Researchers postulated that mirror neurons represent a fundamental mechanism that allows us to understand other's experiences through our own and, thus, underpins numerous social abilities, including action imitation and understanding, mental state inference, and empathy (Gallese et al., 2004; Iacoboni & Dapretto, 2006). These hypotheses led to a drastic increase in scientific interest, reaching its peak in 2013 when over 300 articles were published on the topic of mirror neurons (Heyes & Catmur, 2022).

Despite the initial hype, empirical findings have not provided support for most of these proposals. Neuroimaging, neurostimulation and patient studies have converged on a role for the mirror neuron network in the ability to copy others' bodily movements, as well as low-level action understanding, such as action discrimination and recognition, but not higher-order social cognitive processes like inferring the goals or intentions behind others' actions (for reviews, see Heyes & Catmur, 2022; Thompson et al., 2019). For example, using multivoxel pattern analyses of functional magnetic resonance imaging (fMRI) data, Wurm and Lingnau (2015) showed that activation patterns in the premotor cortex discriminate between videos depicting either opening and closing of a specific object exemplar, but cannot distinguish between opening and closing of different exemplars or object types. This finding suggests that a key node of the mirror neuron network codes only concrete non-generalizable features of actions and not the abstract goal of the action (e.g., to open). Therefore, empirical evidence suggests that the domain-specific mirror neuron mechanism explains only one piece of the puzzle.

This example illustrates that researchers have high hopes for the explanatory power of domain-specific mechanisms alone. Nonetheless, the importance of developing and empirically testing research frameworks that have the potential to explain a broader range of phenomena, regardless of how limited it turns out to be, should not be underappreciated. Now that we know mirror neurons cannot fully account for complex social abilities, we can search for answers somewhere else. The question that becomes relevant is: where should we look for answers? In the next sub-section, I reference the literature on mental state inference to argue that domain-specific frameworks might not be an effective starting point.

1.2.   The theory of mind module

The ability to attribute mental states (e.g., beliefs, desires, intentions) to others, or theory of mind (ToM), has been considered a cornerstone of social interaction that enables us to predict and understand the actions of others (Baron-Cohen et al., 1985; Premack & Woodruff, 1978). Thus, the purported mechanism(s) underpinning ToM ability are the hallmark of most neurocognitive accounts of social cognition (e.g., Adolphs, 2009; Frith, 2007; Happé et al., 2017; Lieberman, 2007). Myriad functional neuroimaging studies have found that, compared to (non-social) inferences about the physical world, inferences about mental states are associated with elevated BOLD responses in a distributed set of brain regions (Figure 1B), including the medial prefrontal cortex, posterior cingulate cortex/precuneus, lateral temporal cortices and the 'temporo-parietal junction' (TPJ) (Molenberghs et al., 2016). There is considerable similarity between the neural correlates of ToM, and those involved in social

cognition more generally (Figure 1) as these same regions seem to also be sensitive to experimental conditions requiring understanding other's actions (Libero et al., 2014; Spunt et al., 2010), sharing other's emotional states (empathy) (Bzdok et al., 2012; Powell et al., 2017; Schulte-Rüther et al., 2007), moral decision-making (Bzdok et al., 2012; Young et al., 2011; Young & Saxe, 2009) and playing strategic games against an opponent (Hampton et al., 2008; Powell et al., 2017; Rilling et al., 2004). All these abilities are thought to rely, to some extent, on inferences about others' affect and/or intentions, indicative of a central role for ToM in social cognition. For example, the network recruited by ToM is thought to support aspects of our ability to understand others' actions that cannot be explained by mirror neuron activity; specifically, while the mirror neuron network represents concrete features of actions, the network recruited by ToM represents the motives behind the actions (Spunt & Lieberman, 2012; Thioux et al., 2008). Therefore, it has been argued that mirror neuron and mentalizing mechanisms have complementary roles in social cognition (Alcalá-López et al., 2019; de Lange et al., 2008; Spunt & Lieberman, 2012; Van Overwalle & Baetens, 2009).

Popular accounts postulate that ToM is underpinned by a singular domain-specific mechanism, such as the meta-representation of mental states in the form of propositional attitudes (e.g., "Sally believes that the ball is hidden inside the box"; Leslie, 1994; Leslie et al., 2004), and the decoupling of others' beliefs from knowledge about reality (Saxe & Kanwisher, 2003; Wimmer & Perner, 1983). Evidence in support of this domain-specific view comes from fMRI studies showing selective activation of the right TPJ when participants are asked to reason about other's false beliefs about the state of reality, but not when reasoning about temporal changes in the state of the physical world (e.g., false photograph) or inaccurate representations of affairs in the physical world (e.g., false map/sign) (e.g., Perner & Leekam, 2008; Saxe & Kanwisher, 2003). Such findings have been interpreted as evidence that the right TPJ is a cortical module specialized for ToM (Saxe & Kanwisher, 2004).

However, this interpretation is not straightforward because the TPJ is a vague term that refers to a functionally and structurally heterogeneous patch of cortex (Mars et al., 2012). The label 'TPJ' refers to a brain area that does not map clearly to anatomical structures and has been used inconsistently to refer to parts of the posterior superior/middle temporal gyrus, angular gyrus and/or supramarginal gyrus, making it difficult for researchers to compare findings across studies and agree on where purported function is localised (Schurz et al., 2017). Moreover, numerous fMRI studies have implicated the right TPJ in a variety of cognitive abilities that do not require mental state inference, including the perception of biological motion, attention reorienting, autobiographical memory, and navigation (Decety & Lamm,

2007; Lee & McCarthy, 2016). Further, disruption of TPJ activity by means of inhibitory transcranial magnetic stimulation (TMS) affects not only ToM (Filmer et al., 2019; Young et al., 2010), but also attention in non-social contexts (Chang et al., 2013; Krall et al., 2016; Lega et al., 2020; Ortiz-Tudela et al., 2018). On this basis, it has been suggested that the TPJ has a domain-general role in attention (Cabeza et al., 2012; Mitchell, 2008), or contextual updating (Geng & Vossel, 2013), and also that it represents a nexus between distinct domain-general cognitive systems (Carter & Huettel, 2013; Decety & Lamm, 2007). While some individual fMRI studies have found that ToM and attention tasks are associated with activation in distinct portions of the right TPJ (Gobbini et al., 2007; Scholz et al., 2009), others reported overlapping activation and successful cross-task classification based on activation patterns, indicative of a domain-general role (Lee & McCarthy, 2016; Mitchell, 2008). It seems that, despite a decade long debate, no consensus has been reached regarding the nature (i.e., domain-specific or domain-general) of the role played by TPJ in ToM.

The case of ToM illustrates how the search for domain-specific explanations of social cognition has, in over 20 years, failed to generate a definitive understanding of the neurocognitive systems underpinning it. Currently, there are multiple possibilities; for example: (i) ToM might rely on a singular specialized mechanism implemented in the right TPJ (e.g., Saxe & Kanwisher, 2003); (ii) ToM could be underpinned by a domain-general function implemented in the right TPJ (e.g., Mitchell, 2008), or (iii) ToM could arise from the interaction of domain-specific and domain-general mechanisms supported by distinct regions within the right TPJ (e.g., Igelström et al., 2016).

Going beyond the contribution of the right TPJ, ToM ability likely results from the interaction of multiple cognitive processes implemented in distinct brain regions/networks, not all of which need to be domain-specific (Gerrans & Stone, 2008; Heyes, 2014; Schaafsma et al., 2015; Schuwerk et al., 2017). This view is supported by individual fMRI studies and meta-analyses that compared ToM tasks differing in terms of paradigm (e.g., false belief stories vs. social animations), stimuli (e.g., verbal vs. non-verbal) and/or instructions (e.g., explicit vs implicit). These studies demonstrated that different task types recruit different brain regions in addition to a core set of cortical areas that includes the right TPJ (Balgova et al., 2022; Gobbini et al., 2007; Molenberghs et al., 2016; Schurz et al., 2014; Spunt & Adolphs, 2014). These findings are in line with the view that ToM ability is underpinned by a set of dissociable underlying processes that are drawn upon differentially depending on task characteristics (Schurz et al., 2020). However, it is currently unclear what these contributing cognitive processes are, how they interact in the service of ToM, and under what circumstances their

contributions are required. In the next section, I argue that generalized frameworks adopted from other established research programmes can provide useful clues.
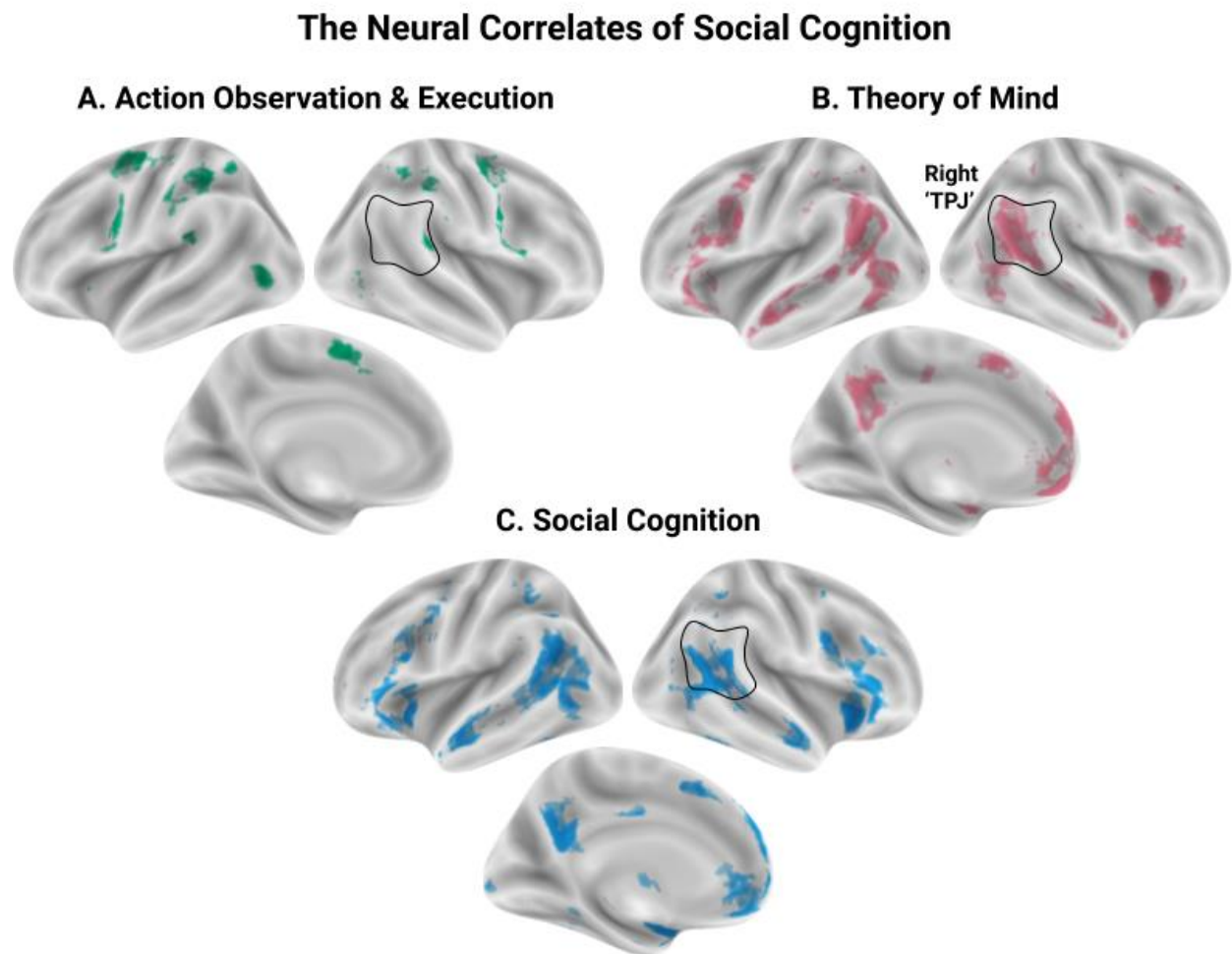
Figure 1. Illustrates brain activation commonly found in response to social cognitive tasks. The maps were generated using Neurosynth automated meta-analysis of functional neuroimaging studies featuring the terms: A) 'action observation' and 'motor performance' to identify the set of brain regions with mirroring properties (see Section 1.1); B) 'mentalizing' and C) 'social cognition'. The right temporo-parietal junction is highlighted - this region has been proposed to act as a functional module for theory of mind and has thus received an elevated status in the social neuroscientific literature. Comparisons between these maps show that (i) the mirror neuron and theory of mind networks are dissociable and (ii) the neural correlates of theory of mind are highly similar to those associated with social cognition more generally.

2. Social cognition through a domain-general lens

The domain-general approach favours a focus on the similarities, rather than the differences, between information processing in what are typically considered social and non-social situations (Amodio, 2019; Binney & Ramsey, 2020; Ramsey & Ward, 2020; Zaki, 2013). A direct implication of such generalist positions is that social and non-social cognition should be underpinned by largely similar neurocognitive systems, and this prediction can be tested empirically. It is important to note that a contribution of domain-general systems would by no means take away from the importance of a sense of 'socialness' or the 'specialness' of our everyday social experiences. Social cognition is at the heart of some of the most rewarding experiences, including social bonding, cooperative endeavours, and human civilization. This is undoubtedly special. Rather, generalist perspectives might provide a fresh starting point for investigations into social cognition that can help disentangle its factor structure, and bring to light the true uniqueness of it. In the following sub-sections, I expand on some of the reasons why studying social cognition through a domain-general lens holds the potential to drive considerable theoretical and empirical progress.

2.1. Completing an incomplete picture

A complete account of social cognition can only be developed if the relationship between domain-specific and domain-general systems is more firmly established. Even if it becomes clear that social cognition relies partly on specialized mechanisms, such as seems to be the case for person perception (Downing et al., 2006) and action understanding (Thompson et al., 2019), a comprehensive model must explain how these interact with general-purpose systems to create a coherent mental representation of the physical world. Indeed, our social interactions are embedded within rich contexts that encompass social, as well as non-social signals, such as inanimate objects, spatial scenes, internal states and motivations that nonetheless affect social behaviour. There is certainly something coherent in the way we experience the co-occurrence of social and non-social cues, indicative of integration between potentially distinct processing streams (Zaki, 2013).

For instance, regardless of whether the right TPJ implements a specialized ToM mechanism, it is clear that its contribution alone is not enough for successful ToM performance. This is supported by neuropsychological investigations showing that intact TPJ in the presence of damage to frontal structures causes impaired ToM ability, and that the nature of the impairment differs depending on lesion location (Rowe et al., 2001; Samson et al., 2005, 2015;

Stone et al., 1998; Stuss et al., 2001). Yet, popular neurobiological models of social cognition mention the domain-general contribution of lateral prefrontal cortices only in passing. For example, Adolphs (2010) pinpoints social regulation as a key component of social cognition that is likely underpinned by domain-general processes. Although possible types of regulatory functions are listed, including cognitive control, emotion regulation, monitoring/error correction, self-reflection and deception, no potential underlying brain mechanisms are specified and most of the surrounding discussion focuses on the regulation of emotional behaviours. Similarly, Frith and Frith (2012) ascribe an important role of a supervisory mechanism in social cognition, which underpins top-down biasing of competition between low-level processing streams and is implemented in the dorsolateral prefrontal cortex and the anterior cingulate cortex. However, the level of description provided is not conducive to testable predictions that can advance research and theory. This is unfortunate, because the cognitive architecture and brain bases of social cognition cannot be elucidated unless the nature of domain-general contributions is understood.

2.2.   Providing a unifying framework

Given that domain-general systems operate across contexts, generalist frameworks can facilitate the development of unifying models that can explain a diverse range of social (and non-social) phenomena. This is important because social cognition is a very broad construct, referring to all mental processes that underpin our ability to interact with conspecifics. This broad construct has been operationalised by breaking it down into a multitude of social abilities and behaviours that have been studied largely in parallel. Particular attention has so far been paid to action understanding and imitation, person and emotion recognition, mental state and trait attribution, empathy, and social decision-making. However, the relationships between these constructs, and the degree to which they rely on common or separable processes, is not clear (e.g., Happé et al., 2017).

The difficulty in delineating the neurocognitive mechanisms underpinning social cognition is partly related to the heterogeneity and inconsistency in the terminology and measures used (Alcalá-López et al., 2019; Cuff et al., 2016; Quesque & Rossetti, 2020). Several terms are often used to describe one construct. For example, the ability to infer other people's mental states has been referred to as theory of mind, mentalizing, mindreading, cognitive empathy, and perspective-taking. Moreover, a single term can have multiple meanings. For example, Cuff et al. (2016) identified 43 discrete definitions of empathy and Thompson et al.

(2019) identified three popular conceptualizations of action understanding, which led to different conclusions about its neural correlates.

Similar issues affect the cognitive measures employed by researchers. For example, the popular reading the mind in the eyes task, which requires inferring other's mental states from pictures of their eyes, was designed as a measure of ToM (Baron-Cohen et al., 2001). However, the same task has also been used to quantify empathy (e.g., Chapman et al., 2006), and emotion recognition ability (e.g., Maurage et al., 2011). Moreover, numerous measures of theory of mind exist and are frequently used (for reviews, see Achim et al., 2013; Quesque & Rossetti, 2020; Schurz et al., 2020), but researchers disagree with regard to their construct validity (Kittel et al., 2022; Olderbak et al., 2019; Quesque & Rossetti, 2020). These issues illustrate how social cognition has been operationalised in various ways in the absence of clear distinctions between different constructs, potentially hampering theoretical and empirical advances.

Instead of encouraging the proliferation of specific social contexts worthy of studying, it might be useful to re-conceptualize social cognition by identifying the similarities between distinct social phenomena. For example, Adolphs (2010) proposed that various forms of social cognition can be reduced to three principal stages of information processing: social perception, social cognition and social regulation (Figure 2A). Social perception involves the sensory transduction and perceptual processing of social signals, such as faces, bodies, and affective touch. Social cognition encompasses mental processes that allow us to go beyond the sensory input, and make inferences about unobservable aspects of the perceived social stimuli, "make[ing] us see the world imbued with social meaning" (Adolphs, 2010). Social regulation has a supervisory role, enabling the control of thoughts and behaviours in order to navigate complex, conflicting and/or ambiguous social signals. In this view, a central aspect of social cognition is ascribed to interferential processing about abstract, unobservable concepts, such as mental states and personality traits. Although the distinction between perceptual, cognitive and regulatory cognition over-simplifies the complex interactions between different levels of information processing, Adolph's (2010) conceptualization highlights a striking similarity between social and semantic cognition. Semantic cognition also refers to the set of cognitive mechanisms that underpin inferences about the unobservable features of sensory inputs and, thus, enables the attribution of meaning to our everyday sensory experiences. From this semantic perspective, the same neurocognitive mechanisms come into play whether we infer that someone is currently tired (ToM) or generally trustworthy (trait attribution), and when we infer that a potted plant needs watering. Therefore, treating social cognition as one way in

which meaning is derived from environmental cues provides an integrative semantic framework that can be equally applied to understanding various forms of social cognition. This example shows how a generalist framework can offer a fresh perspective into the cognitive and brain bases of social cognition by virtue of highlighting conceptual similarities between cognitive domains that have been studied largely in parallel.

The starting point provided by generalist positions contrasts with the commonly adopted domain-specific approach, which assumes that there is something novel that requires additional explanation. Therefore, domain-general frameworks might provide explanations that rely less on computations dedicated to navigating specific social contexts. Of course, such models might be too simplistic and unable to account for the wide range of forms social information processing can take, in which case they can be falsified. Nonetheless, they can help establish in which circumstances our understanding of social cognition can no longer be explained by domain-general mechanisms and additional domain-specific processes are needed. Relatedly, this approach avoids the risk of working to answer questions surrounding social information processing that have already been solved, at least to some extent, in the domain-general literature.

Figure 2. A. Schematic of the main stages of social information processing proposed by Adolphs (2010). B. Schematic illustration of the Controlled Semantic Cognition Framework proposed by Lambon Ralph et al. (2017). Colour coding is used to highlight the correspondence between the two accounts. The putative hub-and-spoke architecture of semantic representation highlights the contribution of both modality-specific cortices distributed throughout the cortex, and of a transmodal semantic hub located in the ventrolateral anterior temporal lobe. The control system is thought to comprise both task-general and semantic components. IFG = inferior frontal gyrus; IFS = inferior frontal sulcus; IFJ = inferior frontal junction; MFG = middle frontal gyrus; IPS = intraparietal sulcus; pMTG = posterior middle temporal gyrus; ATL = anterior temporal lobe.

2.3.   Providing a 'roadmap' for future research

Our knowledge about domain-general systems reflects well-established lines of research that can be effectively leveraged to understand information processing in a social context. For example, our current understanding of semantic cognition is based on a wealth of convergent computational and empirical evidence from neuropsychological, neuromodulation and neuroimaging studies synthesised from across the memory, language and executive function literatures (Jefferies, 2013; Lambon Ralph et al., 2017). This results in a theoretical framework for modelling social cognition that is robust and has firm empirical foundations, and provides novel and falsifiable predictions (Binney & Ramsey, 2020). Specifically, it proposes that common neuro-cognitive resources are deployed for meaning extraction across social and non-social settings (Binney & Ramsey, 2020). As will be discussed in detail below, a main prediction is that social cognition relies on the interaction between a representational system that stores conceptual knowledge and a semantic control system that enables the context-appropriate use of such knowledge, both of which are domain-general. This can be falsified by showing, for example, no or limited overlap between the brain areas that respond to social and non-social semantic tasks. In addition to pointing towards the contribution of a specific set of brain regions and associated functions, the social semantic framework provides clinically-relevant predictions regarding the nature of social impairments caused by damage to these brain regions.

Needless to say, the semantic perspective is not the only generalist framework that can be useful for understanding social information processing. For instance, others have recently highlighted the value of studying social phenomena from the perspective of memory systems (Amodio, 2019), domain-general priority maps (Ramsey & Ward, 2020) and perceptual integration (Zaki, 2013). The rationale for adopting a semantic perspective in this thesis is further motivated in Section 3.3.

2.4.   Interim summary

In sum, applying a domain-general lens to the study of social cognition has a number of strengths: 1) it might allow the generation of a more comprehensive understanding of social information processing by establishing firmer links between social and non-social mechanisms; 2) it might facilitate the development of unifying, generalizable models that can account for multiple forms of social (and non-social) cognition; and 3) it provides a roadmap for social neuroscience research by capitalizing on a wealth of knowledge from established

research programmes to generate falsifiable and clinically-relevant predictions. Given these advantages, this thesis adopts a generalizable framework by treating social cognition as one of the many ways in which we gather meaning form the perceptual world, and addresses outstanding questions arising from this perspective. In doing so, this thesis aims to test whether this semantic approach can expand our understanding of the cognitive and brain bases of social cognition. To this end, insights from one of the most influential accounts of semantic cognition, the Controlled Semantic Cognition (CSC) Framework (Lambon Ralph et al., 2017), will be leveraged. The following section provides an overview of CSC and further motivates the choice of testing whether novel insights into social cognition can be gleaned by adopting a semantic perspective.

3. The controlled semantic cognition framework

3.1. Computational principles

According to the CSC framework, the extraction of meaning from the environment relies on the interaction of two computationally-distinct sub-systems: semantic representation and control (Figure 2B; Jefferies, 2013; Lambon Ralph et al., 2017). The semantic representation system supports the formation and storage of *meaning* or conceptual knowledge, which is acquired through multimodal experience and is deployed to make sense of sensory input. The representation system has a hub-and-spoke cognitive architecture: 'spokes' process modality-specific information and a central supramodal 'hub' integrates this information into coherent, generalizable concepts (for a computational instantiation, see Rogers et al., 2004).

In contrast to *amodal theories* of conceptual representation, which argue that concepts are symbolic and processed independently of the sensorimotor system, the CSC adopts the view from *embodied theories* (for a review, see Meteyard et al., 2012) that modality-specific features, such as visual, auditory and affective information, are the key ingredients for constructing concepts. Moreover, the 'spokes' represent the entry and exist points of the system, supporting the translation from sensation to meaning (for comprehension) and from meaning to motor systems in order to generate meaningful verbal (e.g., speech) and non-verbal behaviours (e.g., use of familiar objects). This view anticipates that the activation of a 'spoke' representation (e.g., the sight of a cup) is able to cue its stored meaning (e.g., object commonly used to drink liquids).

A key tenet of CSC is that interactions between the 'spokes' are mediated by a semantic 'hub'. Computationally, the existence of an intermediary hub solves the challenges associated with the formation of coherent abstract representations (Jackson et al., 2021b; Lambon Ralph

et al., 2007; Rogers et al., 2004). These challenges include the fact that i) the features relevant to a given concept do not always co-occur; instead, they are experienced at different time points and across modalities and contexts (e.g., birds can fly and lay eggs, but not at the same time (Jackson et al., 2021b), and ii) the relationship between conceptual structure and modality-specific features is complex and non-linear (e.g., dalmatians and pugs look different, but they belong to the same conceptual category of 'dogs') (Lambon Ralph et al., 2010). Therefore, simple summation of modality-specific information is not sufficient, and the intermediary hub is crucial for coding the higher-order modality-invariant structure of conceptual knowledge (Figure 3A). This deeper level of conceptual representation can explain the ability to recognize exemplars of the same concept despite differences in their perceptual features (e.g., different breeds of dogs) and to distinguish between exemplars that share perceptual properties but belong to distinct conceptual categories (e.g., tree leaf vs leaf insect).

The second semantic system implements control processes required for retrieving conceptual knowledge that is relevant to current situational demands (for computational models, see Hoffman et al., 2018; Jackson et al., 2021b). Cognitive control processes are necessary because we store more conceptual knowledge about the world than we need in any given situation, and retrieving it all would be inefficient and might interfere with our goals. For example, a pianist does not require to retrieve their expert knowledge about how to play the piano if their only task is to move the instrument across the stage (Saffran, 2000). Therefore, semantic control is crucial for the flexible, context- and task- appropriate use of conceptual knowledge. The control system is thought to include task-general control processes, like working memory and inhibitory control, as well as more specialized forms of control that operate only in semantic contexts (Figure 3B), such as when selecting among competing semantic representations (e.g., 'bank' as in river bank or financial bank) or retrieving weak or noncanonical conceptual associations (e.g., the association between salt-grain, compared to salt-pepper; see Chapter 4 for a more extensive discussion of semantic control mechanisms).

Figure 3. Example cognitive abilities thought to depend on the semantic hub and semantic control mechanisms. A. The semantic hub is considered crucial for coding the higher-order modality-invariant semantic structure. This deeper level of semantic representation explains the ability to (i) distinguish between perceptually similar items that belong to distinct taxonomic categories – e.g., a leaf and a brimstone butterfly whose green wings with strong veining closely resemble leaves; and (ii) recognize conceptual exemplars as belonging to the same taxonomic group even when they share few sensory features – e.g., dahlia flower and tree leaf. B. Semantic control mechanisms are considered crucial for (i) retrieving weak associations – e.g., salt-grain vs salt-pepper, and (ii) selecting among competing alternatives – e.g., select *slice* as the synonym for *piece* when the alternative option is also strongly-associated (*cake* as in the expression *piece of cake*) compared to when the alternative is weakly-associated (*chair*). The correct answers in the examples provided are underlined.

3.2. Neural bases

3.2.1. The 'spokes'

Modality-specific conceptual features are represented in the same unimodal cortices recruited during actual perception and action – the 'spokes' (Figure 2B). A wealth of fMRI studies have found that conceptual retrieval activates modality-specific sensorimotor cortices (Kiefer & Pulvermüller, 2012; Kuhnke et al., 2022). Moreover, the CSC acknowledges that association cortices that process multimodal information, such as affective, language, spatial/location and action systems, can also act as 'spokes'. Importantly, not all 'spokes' contribute equally to all concepts and tasks; rather, the 'spokes' are recruited dynamically depending on the task and type of concept (Chiou et al., 2018; Kuhnke et al., 2020; Pobric et al., 2010b). For example, some concepts are more strongly associated with one modality than another - you can hear the *wind*, but you cannot see it, so the representation of wind would rely more on the auditory 'spoke'. Other concepts, like *liberty* and *infinite,* lack sensorimotor features, so they might rely more on emotion and language 'spokes' (Borghi et al., 2019; Kousta et al., 2011).

3.2.2. The 'hub'

The semantic hub is thought to be underpinned by the bilateral ventrolateral anterior temporal lobes (Figure 2B). This is partly based on evidence from neuropsychological investigations of semantic dementia (SD), a neurodegenerative disorder characterized by atrophy and hypometabolism of the bilateral anterior temporal lobe, which leads to severe semantic impairments despite otherwise relatively preserved cognitive abilities, like visuospatial processing, problem-solving and memory for recent events (Hodges & Patterson, 2007a; Patterson et al., 2007). The semantic impairments affect virtually all conceptual categories (e.g., living and artefacts), input modalities (e.g., written and spoken language, pictures, sounds), and output modalities (e.g., speech, object use), suggestive of a generalized semantic disorder (Bozeat et al., 2000; Coccia et al., 2004; Hodges & Patterson, 2007b; Lambon Ralph & Patterson, 2008; Luzzi et al., 2007; Piwnica Worms et al., 2010). Moreover, when asked to categorise objects, SD patients are influenced by perceptual similarities and consequently over- and under-generalise the category label to incorrect exemplars (cf., Figure 3A). This is consistent with degradation of modality-invariant conceptual information as a result of ATL damage and a consequent over-reliance on perceptual features (Lambon Ralph et al., 2010).

Further support that the ventrolateral ATL functions as a semantic hub comes from cortical stimulation (Shimotake et al., 2015) and functional neuroimaging findings showing

that its ventrolateral portion responds to semantic tasks irrespective of input/output modality or conceptual category (Binney et al., 2010, 2016; Hoffman et al., 2015; Visser et al., 2012; Visser & Lambon Ralph, 2011). Moreover, multivoxel pattern analyses and electrocorticography investigations have showed that the semantic representations in this region are abstracted away from perceptual features (Chen et al., 2016; Coutanche & Thompson-Schill, 2015; Peelen & Caramazza, 2012). The causal contribution of the ATL is supported by studies showing that inhibitory TMS over the lateral ATL induces a domain-general semantic impairment (Pobric et al., 2009, 2010a, 2010b); by contrast, disruption of more posterior regions has modality-selective effects (Pobric et al., 2010b).

3.2.3. Semantic control

Semantic control relies on a distributed left-lateralized set of brain regions (Figure 2B), including the ventrolateral prefrontal cortex, pre-supplementary motor area, intraparietal sulcus and posterior temporal cortex (Jackson, 2021; Noonan et al., 2013). Damage to these regions following left-hemisphere stroke often results in semantic aphasia (SA), an acquired disorder characterized by multimodal semantic impairments. In contrast to SD patients, their semantic deficits are qualitatively different and are suggestive of deficient control and a relatively intact semantic store (Corbett et al., 2009; Jefferies & Lambon Ralph, 2006; but see Chapman et al., 2020). For example, when performing picture naming tasks, SA patients show strong effects of cuing and miscuing, with improved performance in the presence of helpful cues that are consistent with the representation to be retrieved (retrieving tiger when hearing the phoneme "T") and difficulty in the presence of distractors that activate irrelevant conceptual features that must be ignored (retrieving tiger when hearing "L") (Corbett et al., 2011; Jefferies et al., 2008; Noonan et al., 2010; Soni et al., 2009). Further support for a role of these regions in semantic control comes from fMRI studies that have showed they activate more when weak (relative to strong) semantic associations must be retrieved or when there is a need to select one of multiple competing semantic representations (see Figure 3B for example cognitive manipulations) (Badre et al., 2005; Bedny et al., 2008; Snyder et al., 2011; Thompson-Schill et al., 1997; Wagner et al., 2001). Moreover, neuromodulation studies provide evidence for the causal involvement of these regions in semantic control (Davey et al., 2016; Hoffman et al., 2010; Whitney et al., 2011, 2012).

The semantic control network (SCN) partially overlaps with the 'multiple-demand' network (MDN), which is thought to implement task-general executive functions (Duncan, 2010). The areas of overlap include the inferior frontal sulcus/junction, precentral gyrus, pre-

supplementary motor area, intraparietal sulcus and posterior inferior temporal gyrus (Hodgson et al., 2021; Jackson, 2021). An established body of research suggests that the MDN responds reliably to difficulty manipulations across a range of task demands, including working memory (e.g., n-back task), inhibitory control (e.g., Stroop task), sustained attention and vigilance (e.g., stimulus detection, oddball discrimination), cognitive flexibility (e.g., task switching), and planning (e.g., tower of London task) (Camilleri et al., 2018; Fedorenko et al., 2013; Hugdahl et al., 2015; Niendam et al., 2012). This type of control is also relevant when conceptual knowledge is used to guide goal-driven behaviour (Gonzalez Alam et al., 2018). However, functional neuroimaging studies have showed that the SCN and MDN are dissociable, with the SCN additionally recruiting the anterior IFG and posterior middle temporal gyrus (Jackson, 2021). This dissociation is also corroborated by causal evidence showing that inhibitory TMS over the intraparietal sulcus, which is part of the MDN, disrupts both semantic and non-semantic tasks, whereas stimulation of the posterior middle temporal gyrus and left IFG affects only semantic performance (Whitney et al., 2011, 2012). Moreover, activity patterns in the MDN can successfully cross-classify the difficulty of semantic and non-semantic tasks, whereas those in the SCN cannot, indicative of a lack of task-general neural coding of cognitive demands in the SCN (Gao et al., 2021).

3.3.   Towards a general semantics model of controlled social cognition

The rationale for adopting a semantic framework in this thesis is three-fold (also see Binney & Ramsey, 2020). First is the recognition that social interactions often involve extracting meaning from environmental cues. There is no reason for an a priori assumption that the way we derive meaning in social contexts is fundamentally different from non-social contexts. Such a position would be well-justified only if or when evidence accumulates against the involvement of domain-general semantic processes in social cognition.

Second, there is clear correspondence between the three stages of social information processing outlined by Adolphs (2010) and the main components of CSC (Figure 2). Similar to the 'spokes', social perception is concerned with processing perceptual inputs. According to the CSC , going beyond observable sensory inputs and making inferences about unobservable characteristics, an ability attributed to social cognition, depends on access to modality-invariant conceptual knowledge extracted by the semantic hub. Lastly, social regulation, like semantic control, is essential for dealing with ambiguous and conflicting information.

Third, the constructs of conceptual knowledge and cognitive control are not alien to the social neuroscientific literature. Indeed, the importance of social knowledge acquired through

our interactions with others, such as stereotypes and situational scripts, has been recognized (Freeman & Johnson, 2016; Frith, 2007). As such, some researchers have argued for a key contribution of declarative memory systems (Amodio, 2019; Binney & Ramsey, 2020; Spreng, 2013). For example, Kilner (2011) argued that action understanding depends on a semantic retrieval pathway supported by the left IFG and the posterior middle temporal gyrus. Similarly, the control component of CSC mirrors the debate about the necessity of controlled processing in social cognition, which often takes the form of dual-route models that propose the existence of distinct systems for automatic and controlled cognition (e.g., Happé et al., 2017; Satpute & Lieberman, 2006). For example, failure of controlled processing has been invoked to account for anomalous social phenomena such as stereotyping, prejudice and discrimination (Devine, 1989; Fiske, 1998). In addition to these previous accounts, the CSC formalizes the links between the systems for conceptual knowledge and control, providing a mechanistic account of their complementary functions and their interaction, and implicates precise brain structures.

4. Outstanding questions and research aims

The social semantics framework prompts novel questions about the cognitive and brain bases of social cognition, which this thesis begins to answer. In keeping with the key distinction between conceptual knowledge and cognitive control advocated by the CSC, questions surrounding these two systems are addressed separately. Chapters 2 and 3 are concerned with the role of conceptual knowledge, whereas Chapters 4 and 5 are related to the role of cognitive control. The remainder of the present Chapter provides further background and rationale for each subsequent empirical chapter and their research objectives.

4.1. The role of conceptual knowledge: are social concepts special?

According to the social semantics framework, social cognition relies on semantic mechanisms to access and use the wealth of socially-relevant conceptual knowledge we acquire throughout life to guide context-appropriate social inferences and behaviours (Binney & Ramsey, 2020). Therefore, a direct prediction is that the ATL semantic hub should support access to coherent concepts in social contexts. Indeed, there is growing evidence suggesting not only that the ATL makes an important contribution to social cognition, but also that its role seems to be semantic in nature (for reviews, see Binney & Ramsey, 2020; Olson et al., 2013). For example, studies concerned with how knowledge about familiar others is retrieved have identified the ATL as a key contributor that codes identity representations that generalize across perceptual and affective characteristics (Anzellotti et al., 2014; Anzellotti & Caramazza, 2016; Nestor et al.,

2011), as well as the medium through which they are accessed (e.g., via another's name or a picture of their house) (Wang et al., 2017). The findings from this line of work suggest that the person-identity information represented in the ATL is abstract in nature, in line with its role as a semantic hub.

Of particular relevance to the current work is an fMRI study that set out to directly test the prediction arising from the social semantic framework that conceptual processing and high-level social abilities rely on common activation within the ATL (Balgova et al., 2022). Participants were scanned while performing a non-social conceptual similarity judgement task, as well as three ToM tasks. The same region of the left ventrolateral ATL responded to demands for both semantic judgements and mental state inferences, perhaps indicative of shared cognitive processes. Also, this region activated across three different tasks designed to probe ToM and differing in terms of stimuli (verbal, non-verbal) and instructions (explicit, implicit), suggesting that the contribution of the ATL is attributable to the requirement for mental state inference rather than idiosyncratic task features. Interestingly, the magnitude of BOLD change in the bilateral ventral ATL was comparable to that in the right TPJ, which has been ascribed a crucial role in ToM. Overall, this study suggests that semantic and social cognition share neurocognitive correlates in the left ATL, consistent with the view that conceptual knowledge plays an important role in social cognition.

Abstract[2] social concepts, a type of concept that cannot be experienced directly through the senses, are also an important kind of socially-relevant knowledge. To make sense of our social interactions, we often use concepts like those referring to mental (*confusion*) and emotional states (*excitement*), dispositional traits (*honesty*) and types of interpersonal relationships (*friendship*). Such concepts are drawn upon when interpreting complex social signals like other's emotional facial expressions (e.g., Brooks & Freeman, 2018; Gendron et al., 2012; Nook et al., 2015). For example, a pre-requisite for the ability to infer that someone is happy is knowing what happiness means, including how and when it typically manifests (cf.,

---

[2] Here I refer to the construct of *abstractness*, which is distinguishable from *abstraction* (for more extended discussions on this distinction, see Borghi et al., 2019; Dove, 2016). *Abstractness* is a characteristic of concepts whose referents are not single bounded objects/entities that can be directly experienced through the senses. This property distinguishes abstract concepts like *freedom*, *infinity*, and *friendship* from concrete concepts like *animal*, *chair* and *tool*. In contrast, *abstraction* refers to a process in which representation become increasingly decoupled from perceptual information and leads to the emergence of hierarchically ordered categories via generalization over individual exemplars. For instance, the concept *cat* generalizes across specific individual cats, whereas the superordinate category *animals* generalizes over *cats*, *dogs* and other sub-categories. From this perspective, the concept *animal* is more abstract (or higher-order) than *cat*, but its referents are nonetheless perceivable.

Barrett et al., 2011). The representation of social abstract concepts has been explored in two parallel literatures that focused on how meaning is accessed through the medium of language. **Chapter 2** synthesises the insights gained in these two lines of research and identifies important methodological limitations and outstanding questions.

As detailed in Chapter 2, some researchers have claimed that social concepts are stored separately from non-social conceptual knowledge on the basis of observed differences between the neural correlates of social and non-social words (Ross & Olson, 2010). However, others have reported common activation of the semantic hub regardless of the social or non-social nature of the words and argued against these domain-specific claims (Binney et al., 2016; Rice et al., 2018). Yet, comparing the findings of previous studies is challenging because there is no clear distinction between which concepts are considered social and which are not. Concepts' social quality, or *socialness*, has been defined and measured inconsistently across extant behavioural and neuroimaging studies, hampering our ability to draw strong conclusions about socialness as a distinct aspect of semantic representation.

The research reported in **Chapter 3** aimed to address this methodological limitation by collecting and validating socialness ratings for over 8,000 individual words. These ratings were acquired using a novel definition of socialness, which incorporates various previous conceptualisations of the construct. The resulting ratings were used to address two outstanding questions about the role of socialness in semantic representation. First, it is unclear whether socialness captures a distinct aspect of meaning, or merely the same information as other established semantic dimensions, like those quantifying emotional experiences and abstractness. Related investigations have been limited in scale until now, raising questions about the generalizability of findings to larger sets of concepts and word types. Second, to demonstrate that socialness makes an important contribution to the representation of concepts, its association with behavioural indices of conceptual processing should be established. Yet, the relationship between words' degree of social content and lexical-semantic performance has never been assessed.

4.2. The role of cognitive control

4.2.1. Does the semantic control system contribute to social cognition?

There is no doubt that cognitive control mechanisms are crucial for flexible and context-appropriate social and non-social behaviour (Ramsey & Ward, 2020). However, as previously discussed, the contribution of executive functions is underspecified in popular accounts of social cognition. When control has been incorporated into social cognitive models, it was often

from a domain-specific perspective. For example, on the basis of investigations into automatic imitation, it has been argued that social control is underpinned by a specialized self-other distinction mechanism, implemented in the right TPJ, that enables us to distinguish between the representations of our own actions, precepts, and mental states and those of others (Brass et al., 2009; Lamm et al., 2016; Quesque & Brass, 2019; Sowden & Shah, 2014; Steinbeis, 2016). Given evidence that processing these attributes in others recruits the same neural regions as processing one's own experience of them (e.g., mirror neurons) (Brass & Heyes, 2005; Decety & Grèzes, 2006; Lamm et al., 2016), separating self and other-related representations is considered crucial for a number of socio-cognitive abilities, including the suppression of automatic imitation tendencies, visual perspective-taking, ToM and empathy (de Guzman et al., 2016; Happé et al., 2017; Lamm et al., 2016; Santiesteban et al., 2012). However, recent findings resulting from studies with greater statistical power suggest that, at least in the case of automatic imitation, social control is underpinned by domain-general control systems, rather than the TPJ (Darda et al., 2018; Darda & Ramsey, 2019).

Indeed, accumulating evidence points towards an important role for domain-general control and the MDN in some aspects of social cognition. The MDN comprises the precentral gyrus, inferior frontal sulcus extending into the neighbouring inferior and middle frontal gyri, anterior insula, (pre)-supplementary motor area extending into anterior cingulate cortex, intraparietal sulcus and posterior inferior temporal gyrus (Assem et al., 2020; Duncan, 2010). Brain regions that are part of the MDN have been implicated in working memory for social information (e.g., Meyer et al., 2012, 2015; Smith et al., 2017; Xin & Lei, 2015; for a review, see Meyer & Lieberman, 2012), false belief reasoning and self-perspective inhibition (Hartwright et al., 2015; Rothmayr et al., 2011; Saxe et al., 2006; Van der Meer et al., 2011), visual perspective-taking (Bukowski, 2018; Qureshi et al., 2020; Ramsey et al., 2013) and resolving conflicts between social cues (Wang et al., 2022; Zaki et al., 2010). Overall, these findings suggests that domain-general control processes implemented in the MDN, such as interference resolution and inhibitory control, support information processing in social contexts.

According to the social semantics framework, semantic control processes should also be needed to retrieve context-appropriate conceptual knowledge during social tasks (Binney & Ramsey, 2020). A clear prediction arising from this view is that social cognition should rely on the semantic control network, which is partially dissociable from the MDN (Chiou et al., 2022; Gao et al., 2021; Whitney et al., 2011, 2012). To date, this has not been directly tested, although a few studies provide indirect support. Specifically, one study demonstrated that

patients diagnosed with SA, a disorder thought to affect semantic control, show deficient performance on social tasks involving emotion recognition (Souter et al., 2021). This finding is consistent with the idea that semantic control is required to retrieve appropriate emotion concepts. Moreover, an fMRI study has shown that the controlled retrieval and selection of social concepts activates the left IFG, which is a key node of the SCN (Satpute et al., 2014). However, it is unclear whether the SCN contributes to higher-order social abilities, such as mental state inference.

The main aim of **Chapter 4** is to directly test whether regions that respond to semantic control demands are recruited by social tasks. To this end, the neural correlates of social cognition and semantic control will be compared. The study design addresses two methodological issues discussed above. First, social cognition is a multifaceted construct encompassing a range of social abilities. Therefore, we selected four distinct abilities thought to rely on inferential processing and studied their neural correlates in parallel. We predicted that brain regions activated by all these abilities will include key nodes of the SCN. If semantic control regions are engaged by these various forms of social information processing, this would provide empirical support for the claim that the semantic framework offers a unifying account of social cognition. The second consideration was the considerable impact that the choice of paradigm, instructions and stimuli can have on the neural correlates observed in individual fMRI studies (cf., Balgova et al., 2022). Therefore, we adopted a meta-analytic approach that can identify reliable trends in extant functional neuroimaging data.

4.2.2. Does the left inferior frontal gyrus make dissociable contributions?

The study reported in Chapter 4 revealed that the left IFG, a key node of the SCN, is reliably recruited by all four social abilities investigated. This finding is perhaps not surprising given that this region has been implicated in numerous cognitive domains, to the extent that it has been ascribed a general-purpose control function in some accounts (Goghari & MacDonald, 2009, 2009; Hugdahl et al., 2015; Swick et al., 2008; Tops & Boksem, 2011). However, alongside common activation in the posteroventral IFG, the four types of social tasks engaged other portions of the IFG differently. This suggests that, instead of supporting a single domain-general function, the left IFG might have multiple dissociable functions. Indeed, there is evidence to suggest that the left IFG makes dissociable contributions to semantic control. Particularly, while the dorsal and posterior portions of left IFG are responsive to both semantic and non-semantic control demands, in line with a domain-general control function (Assem et al., 2020; Diachek et al., 2020; Fedorenko et al., 2013), the anteroventral left IFG contributes

to challenging semantic tasks specifically (Chiou et al., 2022; Gao et al., 2021; Hodgson et al., 2021). Therefore, it is possible that the left IFG underpins multiple forms of control that are recruited differently during social information processing depending on task/context. To gain insight into the functional organisation of the left IFG, **Chapter 5** provides a detailed data-driven exploration of its voxel-wise patterns of functional connectivity. In line with the domain-general approach, the left IFG is studied in the context of its cross-domain contributions, enabling insight into its core function(s), rather than as restricted to social contexts.

The final chapter, **Chapter 6**, summarises the findings from Chapters 3-5 and highlights their implications for our understanding of the neurocognitive underpinnings of social cognition. The chapter concludes by reflecting on whether studying social information processing from a semantic perspective can be a fruitful endeavour.

CHAPTER 2

Social semantics: a role for socialness in the representation of abstract concepts?

This chapter has been accepted for publication as:

Pexman, P.M., Diveica, V. & Binney, R.J. (accepted). Social semantics: the organization and grounding of abstract concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Authors' contributions:

Veronica Diveica surveyed the literature to compile the list of definitions presented in Table 1 and collected and analyzed the socialness data illustrated in Figure 1. Veronica Diveica, Dr Richard Binney and Prof Penny Pexman contributed equally to the writing of this paper.

Open science practices and data/code availability statement:

The data illustrated in Figure 1 are available via the Open Science Framework at osf.io/2dqnj.

Abstract

Abstract concepts, like *justice* and *friendship*, are a central feature of our daily lives. Traditionally, abstract concepts are distinguished from other concepts in that they cannot be directly experienced through the senses. As such, they pose a challenge for strongly embodied models of semantic representation that assume a central role for sensorimotor information. There is growing recognition, however, that it is possible for meaning to be 'grounded' via cognitive systems, including those involved in processing language and emotion. In this article, we focus on the specific proposal that *social* significance is a key feature in the representation of some concepts. We begin by reviewing recent evidence in favour of this proposal from the fields of psycholinguistics and neuroimaging. We then discuss the limited extent to which there is consensus about the definition of 'socialness' and propose essential next steps for research in this domain. Taking one such step, we describe preliminary data from an unprecedented large-scale rating study that can help determine how socialness is distinct from other facets of word meaning. We provide a backdrop of contemporary theories regarding semantic representation and social cognition and highlight important predictions for both brain and behaviour.

1. Introduction

> *"You are mistaken, Mr Darcy, if you suppose that the mode of your declaration affected me in any other way, than as it spared me the concern which I might have felt in refusing you, had you behaved in a more gentleman-like manner."*
> *– Jane Austen, Pride and Prejudice (1813/2019)*

This brief extract from *Pride and Prejudice*, a classic tale in the importance of personal character, integrity and morality, is rich with references to concepts of a social nature (e.g., *manner*, *gentleman*, and *refuse*). Indeed, a large portion of even the most everyday vocabulary is occupied by abstract words imbued with a sense of socialness. Arguably, this reflects the vital role of social conceptual knowledge in navigating our interpersonal world. After all, humans are intrinsically and uniquely social. We exhibit a natural propensity to cooperate, to coordinate, and to learn from one another, and to a very large extent this is done though the medium of language. It is argued that our advanced social cognitive and emotional abilities, as well as the evolution of language, are an adaptation to, and thus a direct consequence of life lived in groups (Tomasello, 2009; Tomasello, 2020). By extension, this suggests there could be a fundamental nature to the social qualities of words.

Recent work in the field of cognitive science has begun to elucidate the ways in which socialness impacts the structure of concepts and the representation of semantic knowledge in the human brain, and this work will be the subject of the first two parts of this paper. In Part A, we will begin with a brief overview of general theories of semantic memory, with a particular emphasis on what is known as the grounding problem and the difficulties it poses for representing abstract word knowledge. Then we will introduce nascent theories that posit social experience as a mechanism for grounding conceptual knowledge, together with a review of recent semantic feature generation/ratings studies that identify socialness as an important factor for distinguishing amongst different 'types' of concepts. In Part B, we will review a set of neuroimaging studies that have approached socialness from a different methodological perspective, exploring if and how socialness of concepts is represented at the level of macroscale brain anatomy. This includes evidence that is in line with claims that social concepts have a special, or even privileged status over other types of concepts, and suggests socialness drives the functional organisation of neurobiological systems.

Moreover, a key aim of this paper is to highlight major outstanding questions, and this includes one very fundamental issue that arises from the work described in both Parts A and B; what is it exactly that defines a word as being 'social'? In Part C, we will discuss the limited

extent to which there is consensus on the kinds of semantic features that amount to 'socialness' and the degree to which it has been established as a valid and meaningful construct. Consequently, we argue that to further progress theory, the field must first establish a clearer working definition of socialness. To this end, we describe preliminary data from a large-scale rating study in which Diveica et al. (2022) provided participants with an inclusive definition of socialness and asked them to collectively rate over 8,000 English words. This includes findings that appear to confirm that these ratings capture aspects of word meaning that are distinct from those measured via other semantic variables like concreteness.

The issue of whether social concepts are a distinct type, either from other forms of abstract concept, or even more generally speaking (i.e., such that this extends to concrete social concepts), has important theoretical implications regarding the fundamental organisational principles underpinning semantic representation.[1] In turn, it has implications for our understanding of the configuration of brain systems, including those responsible for language and social cognition. These implications extend to applied areas of research where an improved framework for understanding the way social and affective concepts are learned, represented and impaired, could have important implications for educational and clinical practice (see Binney, Zuckerman, et al., 2016; Nook et al., 2020).

2. Part A – Abstract word representation: a role for socialness?

There are now numerous theories of semantic knowledge, which vary in the extent to which sensorimotor information (e.g., visual, auditory, or tactile experience) plays a role in the representation and processing of word meaning. At one end of the spectrum, amodal theories posit that semantic knowledge is represented symbolically, distinct from the ways we experience the world (e.g., Pylyshyn, 1984; Quillian, 1969). At the other end, strongly embodied theories posit that knowledge is represented by sensory and motor systems (e.g., Glenberg, 2015; Glenberg & Gallese, 2012). Between these poles lie multimodal or multiple representation theories, which posit that semantic knowledge is represented in many ways (e.g., via language, emotion, introspection, and sensorimotor systems), and some versions of those theories include an intermediary supramodal hub (e.g., Lambon Ralph et al., 2017; Patterson et al., 2007). The hub accounts position language information as one of many types of knowledge connected to the hub. Further, the multiple representation accounts assume that

---

[1] There is some debate about whether the conceptual system is separate from the lexical-semantic system. We take the position that the conceptual system is not separate from the semantic meaning accessed during language processing.

different kinds of information are important for different types of concepts (e.g., Barsalou, 2008; Borghi et al., 2019).

Proponents of semantic theories that include reliance on sensorimotor systems have argued that these theories have the advantage of addressing the grounding problem (Harnad, 1990; Searle, 1980). The grounding problem asks, in essence, if knowledge is represented as symbols, then how do those symbols map to the world? Embodied theories solve that problem by proposing that cognition engages modal systems (e.g., those used for perception, action) to represent semantic knowledge. Strongly embodied theories, however, run into difficulty explaining representation of abstract words. The meanings of abstract words, by definition, cannot be learned or experienced through sensorimotor systems, so they cannot be accounted for by embodied theories. To explain knowledge of abstract words, other means of learning and representation must be considered. Barsalou and colleagues (e.g., Barsalou & Wiemer-Hastings, 2005; Barsalou, 2020) have noted that too many approaches to abstract concepts emphasize what they do not contain (sensorimotor information) and that a more positive approach is needed to explore what they do contain. To that end, Barsalou and Wiemer-Hastings (2005) used a property generation task to compare the features of a small set of abstract and concrete words (see also Wiemer-Hastings & Xu, 2005). They found that abstract words were notably different in that their meanings were mainly associated with introspections and, in particular, social aspects of situations, such as people, communication, and social institutions.

In addition, work has begun to identify concept "types" within the abstract realm. Much of this work is inspired by multiple representation views, and considers multiple sources of grounding beyond the sensorimotor, including the potential contributions of action, language, interoception, emotion, cognition, and other internal states. Notably, Borghi and colleagues have proposed the Words as Social Tools (WAT) account, which focuses particularly on the acquisition and representation of abstract word meaning (Borghi et al., 2019; Borghi & Binkofski, 2014). They argue that abstract words are associated with richer linguistic, inner and, importantly for present purposes, *social* experience, than are concrete words (also Zdrazilova & Pexman, 2013). Further, they suggest that there could be different types of abstract concepts which vary in their reliance on these different types of information. They suggested that these types of abstract concepts might include institutional, temporal, mental states, emotional, numerical, and social concepts.

In related empirical work, researchers have explored the features or properties of abstract word meanings in order to derive potential clusters or types. For instance, Harpaintner

et al. (2018) examined the features listed for 296 abstract words and found that they fell into three clusters. The largest cluster was primarily distinguished by a higher proportion of sensorimotor features, with some social features. A second cluster was distinguished by a high proportion of internal/emotional features and more social features than either of the other clusters. The third, smallest cluster was distinguished by a high proportion of verbal association features. Similarly, Troche et al. (2014) investigated the organization of abstract and concrete English nouns by asking participants to rate 200 concrete and 200 abstract words on 12 dimensions. They analyzed the ratings and identified three latent semantic factors: affective association/social cognition, perceptual salience, and magnitude (also see Troche et al., 2017). Abstract word meanings relied more heavily on affective association/social cognition than did concrete meanings. Villani et al. (2019) asked participants to rate 425 abstract Italian nouns on 15 dimensions and identified four clusters: philosophical/spiritual concepts, physical, spatio-temporal and quantitative concepts, emotional/inner state concepts, and self and sociality concepts. Additional analyses showed that the involvement of the dimension they called social metacognition (defined as a reliance on other people to understand the meaning) distinguished abstract from concrete words, with more abstract words tending to have higher ratings of social metacognition. In addition, ratings on a dimension that they termed social valence (defined as evocative of social situations) were associated with emotion ratings, and with ratings of mouth movement and hearing. These latter relationships were attributed to the important role that language is assumed to play in representing abstract concepts, and to the importance of mouth movement and hearing to language.

Similar conclusions about the existence of types of abstract concepts were drawn from an fMRI study reported by Vargas and Just (2020). They investigated the clustering of 28 abstract words in terms of neural signatures, after participants were scanned while thinking of properties of each word. Results showed that there tended to be commonalities across participants in terms of the neural signatures of each word, and the authors identified three latent factors including verbal representation, externality/internality, and social content (also see Huth et al., 2012, 2016).

Thus, there is evidence from some property-generation and feature-rating studies that social words may be a distinct type of abstract word, consistent with assumptions of the WAT theory and other multimodal accounts. Each of these studies, however, has involved a relatively small sample of abstract words, many fewer than people actually know. Therefore, it is possible that the results could be specific to the words tested and may not generalize to a larger set. Thus, there is a need to explore socialness at a much larger scale and right along the

concreteness continuum. There is also a need to investigate behavioural effects of socialness. That is, if social words are a distinct type, then one might expect that a word's degree of socialness would be reflected in some way in behavioural measures of lexical-semantic processing, as much as semantic dimensions like valence (Warriner et al., 2013) and concreteness (Brysbaert et al., 2014) are related to such processing (e.g., Kousta et al., 2011; Newcombe et al., 2012; Pexman & Yap, 2018). One might also expect behavioural responses to social abstract words to be different to those given to other types of abstract words (see Muraki et al. (2020) for an example of this approach). And yet, comparisons between social abstract words and other abstract word types have still to be made in the context of larger scale behavioural studies. However, they have been contrasted in the neuroimaging literature, reviewed next.

3. Part B – Socialness and the brain

A review of neuroimaging literature concerning the representation of abstract concepts identified a small number of papers that treat social concepts a priori as a discrete 'category' (Conca et al., 2021). Most of these studies contrasted social words with a more general class of abstract or concrete words and set out to identify common activity, and/or that which is uniquely associated (Binney, Hoffman, et al., 2016; Rice et al., 2018; Ross & Olson, 2010; Zahn et al., 2007). The earliest of these studies generated a hypothesis that social concepts are a class of concepts with a special, or even privileged status over other types of conceptual knowledge (Ross & Olson, 2010; Zahn et al., 2007). In this context, social conceptual knowledge has been broadly defined as person-specific knowledge (Simmons et al., 2010), but also knowledge about interpersonal relationships, social behaviours, and of more abstract social concepts such as *truth* and *liberty* (Olson et al., 2013; Zahn et al., 2007). These early studies revealed a patch of anterior temporal association cortex that the authors claimed is selectively involved in processing semantic information of a social nature (Olson et al., 2013; Simmons et al., 2010).

The 'social knowledge hypothesis' (Olson et al., 2013; Zahn et al., 2007) can be likened to other forms of 'multiple semantics' views (Meteyard et al., 2012; Shallice, 1987; Warrington & McCarthy, 1994) in which the semantic system is composed of multiple independent stores that are differentiated by their link to distinct sensorimotor modalities. Of course, the difference is that the social distinction is based on domain-specificity rather than modality. To understand how this hypothesis formed the starting point for this particular set of neuroimaging studies, one can look to the broader social neuroscience field from which they stemmed. The emergence

of this field was triggered, at least in part, by the 'social brain hypothesis' (Brothers, 1990; Dunbar, 1998), which states that the expansion of frontal and temporal neocortices across primate species in the human evolutionary lineage is explained by their increasingly high levels of sociality (for a related review, see Braunsdorf et al., 2021). This created a pervasive assumption, sometimes implicit, that there is a circumscribed set of brain regions that are dedicated to, and, by inference, support specialized processes for social perception and cognition (Brothers, 1990; Frith, 2007). The extent to which domain-specificity of systems for processing social information exists is hotly debated (Amodio, 2019; Binney & Ramsey, 2020; Kilner, 2011; Spunt & Adolphs, 2017), but there is evidence for the existence of brain regions or pathways that are sensitive to socialness, particularly at the level of perceptual processes (Adolphs, 2010; Kanwisher, 2010; Pitcher & Ungerleider, 2021). This includes visual cortex with ostensibly selective engagement by faces (Kanwisher & Yovel, 2006), bodies (Downing et al., 2001), and social interactions (Isik et al., 2017). Whether this putative specialisation cascades downstream to higher-order cognitive systems (e.g., memory; executive function) is a more contentious issue (Adolphs, 2003; Binney & Ramsey, 2020; Mars et al., 2012; Rushworth et al., 2013).

To date, the leading candidate in terms of a locus for a selective social semantic system lies within the dorsolateral aspects of the anterior temporal lobe (ATL), specifically the anterior superior temporal gyri/sulci (Arioli, Gianelli, et al., 2021; Olson et al., 2013; Rice, Lambon Ralph, et al., 2015). These ATL subregions exhibit elevated BOLD responses when semantic judgments made on socially relevant stimuli are compared to those made on non-social stimuli (Binney, Hoffman, et al., 2016; Lin et al., 2019; Lin, Wang, et al., 2018; Rice et al., 2018; Ross & Olson, 2010; Zahn et al., 2007). The dorsolateral ATL also appears to increase its response in line with an accumulation of social meaning across connected text (Zhang et al., 2021). The role of the ATL in representing social knowledge has been ascribed with a right lateralisation within some accounts (Gainotti, 2015), although individual fMRI studies (Binney, Hoffman, et al., 2016; Lin et al., 2015, 2019; Rice, Lambon Ralph, et al., 2015; Ross & Olson, 2010) and meta-analyses (Arioli, Gianelli, et al., 2021; Rice, Lambon Ralph, et al., 2015) indicate bilateral involvement (also see Eleonora Catricalà et al., 2020; Pobric et al., 2016).

More recent neuroimaging studies have attempted to disentangle the socialness effect driving some ATL activations from other potentially confounding variables. For example, it is possible that the social concepts explored in neuroimaging studies are, on average, more abstract than more general classes of concepts. However, studies have shown that preferential left hemisphere dorsolateral/polar ATL activation cannot be easily explained by differences in

concreteness, or at least imageability, between social words and control words (Binney, Hoffman, et al., 2016; Lin et al., 2019; Lin, Wang, et al., 2018; Wang et al., 2019), nor by differing degrees of multiplicity of single word meanings (sometimes referred to as 'semantic diversity'; see Hoffman et al., 2013) (Binney, Hoffman, et al., 2016). A putative involvement of these regions in combinatorial conceptual processes does also not appear to explain differential engagement by social and non-social words (Lin et al., 2020; Lin, Yang, et al., 2018). Many of these studies have also been careful to rule out explanations in terms of fundamental lexical properties such as word frequency or syllable/word length (Binney, Hoffman, et al., 2016; Lin et al., 2015, 2019; Lin, Wang, et al., 2018). Another semantic factor that could covary with socialness, and account for some preferential activations, is emotional valence. Indeed, one study has shown that social-emotional stimuli elicit stronger responses in the left dorsolateral/polar ATL than other social words, which in turn activate the region more strongly than stimuli lacking any social meaning (Mellem et al., 2016). However, Wang et al. (2019) demonstrated at least partially dissociable responses across left lateral ATL subregions to the socialness, valence and abstractness dimensions underlying word meanings. Overall, this collection of neuroimaging studies suggests that the socialness of a concept makes a unique contribution to driving differential recruitment of the brain regions involved in processing meaning (for neuropsychological evidence, see Catricalà et al., 2014; 2021). Moreover, they are to some extent compatible with the claim that there is a semantic system dedicated to the representation of social conceptual knowledge, and that this is located in the left dorsolateral/polar ATL (Olson et al., 2013; Simmons et al., 2010; Zahn et al., 2007).

Of course, an alternative to the notion of a 'social brain' or, more precisely, that there are networks or subsystems specialised for social processes, is that social cognition is underpinned by a set of domain-general systems (Amodio, 2019; Binney & Ramsey, 2020; Spunt & Adolphs, 2017). As alluded to above, from a strong version of this perspective, socialness effects at the levels of brain and behaviour reflect variations amongst more general properties of stimuli and/or task demands, rather than socialness per se. From a more compromising perspective, it is argued that social interaction could draw on an array of neurocognitive systems in something of a unique way, but, fundamentally, those systems are built for more generalised processes (e.g., Binney & Ramsey, 2020; Spunt & Adolphs, 2017). For example, an alternative to domain-specific accounts of ATL function like the social knowledge hypothesis is the 'graded semantic hub' account proposed by Binney, Hoffman, Lambon Ralph and colleagues (Balgova et al., 2021; Binney et al., 2012; Binney, Hoffman, et al., 2016; Rice, Hoffman, et al., 2015). According to this framework, the whole ATL comprises

a unified semantic representational space, all of which is engaged by the encoding and retrieval of concepts, and by concepts of any kind. At the centre of this space lies the ventrolateral ATL, which has a supramodal semantic function, meaning that its engagement during semantic processing is invariant to, for example, idiosyncratic task features, including the modality through which concepts are accessed. Near the edges of this space, however, there are connectivity-driven gradual shifts in semantic function towards subspecialisations for processing certain types of semantic features (for a computational exploration of this general hypothesis, see Plaut, 2002). This might include, for example, at the dorsolateral aspects, a specialisation for processing socio-emotional semantic features (Binney, Hoffman, et al., 2016), which could arise from greater proximity and connectivity to the limbic system (Binney et al., 2012; Jung et al., 2017; Plaut, 2002). Consistent with this account are a series of neuroimaging studies by Binney, Lambon Ralph and colleagues which show that, when care is taken to ensure that fMRI signal can be acquired from across the whole ATL, it becomes clear that the ventrolateral ATL activates strongly and equivalently during semantic judgements made on social and non-social stimuli (Binney, Hoffman, et al., 2016; Rice et al., 2018; also see Balgova et al., 2021). This same ventrolateral ATL region is implicated in general semantic processing in several neuropsychological, neurostimulation, neuroimaging, and electrophysiological studies that have used a variety of verbal and nonverbal tasks/stimuli (Binney et al., 2010; Binney & Lambon Ralph, 2015; Chan et al., 2011; Marinkovic et al., 2003; Pobric et al., 2007, 2010a; Shimotake et al., 2015; Visser & Lambon Ralph, 2011). Critical for this graded hub account is the additional fact that the omni-category response of the ventrolateral ATL is much greater in magnitude than that of the social-selective response of the dorsolateral ATL (Binney, Hoffman, et al., 2016; Rice et al., 2018). Therefore, these latter studies suggest that, at least within the ATL, differences in the way the brain is engaged by social and non-social concepts are small, or subtle, compared to the similarities. This is consistent with the claim that, rather than there being distinct systems for social and general semantic representation, there is a single domain-general conceptual system and parts of this system are dynamically and differentially engaged by different types of meaningful stimuli and semantic task demands (cf., the Social Semantics framework outlined by Binney and Ramsey, 2020; also Lambon Ralph et al., 2017).

The graded hub hypothesis is an extension of the hub-and-spoke model of semantic representation proposed by Patterson, Rogers, Lambon Ralph and colleagues (Lambon Ralph et al., 2017; Patterson et al., 2007). According to this framework, the ATL sits at the heart of a spoked semantic architecture comprised of association regions involved in modality-specific

sensorimotor processing, as well as affective and linguistic processes. The hub-and-spoke model emphasises that semantic representation arises from the *conjoint* action of modal systems and an intermediary supramodal hub (Lambon Ralph et al., 2017). It offers a reconciliation between distributed-only embodied accounts in which concepts are dependent upon systems involved in sensory and motor processing (Barsalou, 2009; Gallese & Lakoff, 2007; Pulvermüller, 2001; Vigliocco et al., 2004) and neuropsychological and computational modelling data that point towards the existence of a hub (e.g., Lambon Ralph et al., 2010; Rogers et al., 2004). A fuller discussion surrounding the necessity of a hub is beyond the scope of this review, and for a starting point we refer the reader to Lambon Ralph et al. (2017), as well as Meteyard et al. (2012). However, we have chosen to raise this broader hub-and-spoke proposal here because it is a neurobiologically-constrained model that, like multimodal or multiple representation views, acknowledges sources of semantic information beyond sensorimotor experience, including contributions from language, emotion, and other internal states (Lambon Ralph et al., 2017). Moreover, like some of the multimodal views described in the previous section (e.g., Barsalou, 2008; Borghi et al., 2019), it hypothesises that different types of concepts (e.g., tools) can vary in their reliance on different sources of information (e.g., object affordances, and kinematics), which will be reflected in differential engagement of spoke regions (Chiou & Lambon Ralph, 2019; Pobric et al., 2010b; also see Wilson-Mendenhall et al., 2013). This notion lends one interpretation to neuroimaging studies that investigate social concept representation and implicate brain regions outside of the ATL. For example, two recent studies have demonstrated an apparent selective engagement of the precuneus, a region associated with visual-spatial imagery (Bzdok et al., 2015), during the processing of abstract social words (Leshinskaya et al., 2017; Vargas & Just, 2020). This could reflect a tendency for social concepts to draw differentially upon systems that capture visual or spatial elements of interpersonal contexts (Vargas & Just, 2020).

4. Part C – What is 'socialness'?

In the sections above we have provided a brief overview of two parallel literatures amongst which socialness has begun to emerge as an important organisational principle underpinning semantic representation. In Part A, we described property generation and feature rating studies that have explored the attributes of abstract words and have extracted socialness as a latent factor that distinguishes abstract from concrete words, and even distinguishes different types of abstract words. In Part B, we reviewed a literature that has emerged in parallel, describing a set of neuroimaging studies that have probed socialness as a predictor of differential patterns

of brain activation evoked during semantic processing. In contrast to property-generation research, most of these neuroimaging studies approached social concepts as an a priori discrete type of concept. This has, thus far, been fruitful in that this brain-based evidence points to socialness being independent of more general semantic properties, such as abstractness, emotional valence and other facets of single word meaning. There is now a burgeoning debate regarding the relative size of the contribution that socialness makes to semantic representation. On one hand, it has been argued that social words are a distinct type and, moreover, that there are specialised neural systems underpinning social semantics. On the other hand, socialness can be framed as one of many dimensions that coexist to define a single representational space underlying general semantics.

However, we assert that, while these lines of research are both intriguing and promising, the conclusions and discussions that have transpired from them are mostly premature, because the ostensive evidence has accumulated in the absence of clear boundaries between what is social, and what is not. This is true both at the level of theory and in the empirical measures. Without agreeing on this definition, at least to some extent, it will not be possible to compare theories and evaluate evidence in support of them. So, what is *socialness* actually?

Socialness as a construct has been characterised variably in terms of behavioural descriptiveness, and social concepts have been distinguished from non-social concepts on divergent sets of criteria. To illustrate this point, we have collected examples in Table 1 (also see Conca et al., 2021). Many of these studies focused on a word's reference to social interaction by measuring, for example, the extent to which a word refers to relationships between people (Troche et al., 2014, 2017), or how often its referent involves interaction between people (Lin et al., 2019, 2020; Lin, Wang, et al., 2018; Wang et al., 2019). In contrast, other definitions emphasize specific aspects of social experience, such as how well a word describes social behaviours (Zahn et al., 2007), or the degree to which word meanings relate to the relationship between self and others (Crutch et al., 2012).

Table 1. Definitions used to measure socially-relevant semantic constructs in previous studies.

Publication	Name of Construct	Type	Definition
Arioli et al. (2021)	sociality	dimension	how much a word inherently refers to information concerning social as opposed to individual contexts
Binder et al. (2016)	social	dimension	the degree to which one thinks of a thing as an activity or event that involves an interaction between people
	human	dimension	the degree to which one thinks of a thing as having human or human-like intentions, plans, or goals
	communication	dimension	the degree to which one thinks of a thing as a thing or action that people use to communicate
	self	dimension	the degree to which one thinks of a thing as related to your own view of yourself, a part of your self-image
Catricalà et al. (2020)	social	dimension	how much a word is linked to a social situation or to an interaction among people, both in terms of inclusion and exclusion
Crutch et al. (2012)	social interaction	dimension	the degree to which concepts relate to the relationships between self and others
Diveica et al. (2022)	socialness	dimension	the degree to which a word's meaning has social relevance by describing or referring to a social characteristic of a person or group of people, a social behaviour or interaction, a social role, a social space, a social institution or system, a social value or ideology, or any other socially-relevant concept
Harpaintner, et al. (2018)	social constellation	category	a feature or a situation that describes the coexistence of different persons or which implies an interaction between at least two different persons
Lin et al. (2015)	sociality	dimension	the number of people involved in an event to which a verb refers
	biological motion	dimension	the extent to which the meaning of a verb brings to mind biological motion

Lin et al. (2019, 2020) Lin, Wang, et al. (2018) Wang et al. (2019)	sociality	dimension	how often the meaning of a word/the use of an object involves an interaction between people
Mellem et al. (2016)	social content	category	referring to people either by a proper name or the name of an occupation/title
Roversi et al. (2013)	institutional objects	category	an artefact that performs its function via the collective acceptance displayed by a given community (status function) and not in virtue of its physical features
	social objects	category	an entity that presupposes the existence of at least two agents engaged in some form of common activity and that does not have a clear status function attached to it
Troche et al. (2014, 2017)	social interaction	dimension	the degree to which the word relates to relationships between people
	morality	dimension	the degree to which the word relates to morality, rules, or anything that governs one's behaviour
Vargas and Just (2020)	social content	dimension	the degree to which the concept involves social interaction or self-perception as affected by social interaction
Villani et al. (2019)	social metacognition	dimension	how much others were needed to understand the meaning of the word
	social valence	dimension	the degree to which the concept evokes social circumstances
Villani et al. (2021)	pure institutional objects	category	entities constituted by formalized rules in a social framework
	meta-institutional objects	category	concepts that are necessary to define the content of institutions but are not defined by those institutions

Zahn et al. (2007)	behaviour descriptiveness	dimension	how well a word describes a detailed specific set of social behaviours of persons
	social category breadth	dimension	how many different kinds of social behaviours of persons a word can apply to
Zhang et al. (2021)	social semantic richness	dimension	the extent to which the word/sentence/narrative is related to interactions between people

Following a review of the material presented in Table 1, we suggest that there are two distinct emerging approaches to the construct of socialness. On one hand, there are social measures designed to capture contextual information, such as the degree to which a word meaning evokes a set of social circumstances (Villani et al., 2019), or whether it applies to social as opposed to individual contexts (Arioli, Basso, et al., 2021). On the other hand, there are measures probing specific social features of word meaning such as the scale of interaction/number of agents implicated (Lin et al., 2015) and the degree to which a referent has human-like intentions, plans, or goals (Binder et al., 2016). This distinction might reflect different representational frameworks for meaning such as those based on features/similarity and those based on association (Crutch et al., 2007; Crutch & Warrington, 2005), and it could be an important avenue for future research into the mechanisms by which socialness is attributed to concepts. However, the heterogeneity in definitions across this set of studies is striking, highlights theoretical inconsistencies, and hinders our ability to compare findings across studies. Certainly, it imposes grave limitations on the conclusions that can be made presently about *socialness* as a neurobiologically and/or behaviourally relevant principle.

We argue that, to further progress theory, the field must first establish a clearer working definition of *socialness*. Further, the field would be advanced if large-scale norms of rated socialness were available, much as they have been made available for thousands of words for semantic variables like concreteness (Brysbaert et al., 2014), emotional valence (Warriner et al., 2013) and others. We believe this can best be achieved, at least initially, by adopting a broad definition of socialness. To aid this endeavour, we recently obtained ratings for 8,388 English words by asking participants to rate socialness according to the following definition (Diveica et al., 2022):

> *the extent to which each word has social relevance by describing or referring*
> *to a social characteristic of a person or group of people (e.g., 'trustworthy'), a*
> *social behaviour or interaction (e.g., 'to fight'), a social role (e.g., 'teacher'),*
> *a social space (e.g., 'pub'), a social institution (e.g., 'hospital') or system (e.g.,*
> *'nation'), a social value (e.g., 'righteousness') or ideology (e.g., 'feminism'), or*
> *any other socially-relevant concept.*

To our knowledge, the resulting norms are the largest set of openly available word socialness ratings. We believe that employing an inclusive definition was a crucial next step for understanding the construct of *socialness*. This allowed us to test the extent to which socialness is reliably perceived as a broad construct, and as applicable to various types of

words/parts of speech. Initial explorations of the ratings reveal that, when broadly defined, socialness ratings have good reliability and validity. We have begun exploring to what extent these new socialness ratings capture aspects of word meaning that are distinct from those measured via other semantic variables, such as concreteness and emotional valence (see Figure 1). Results showed that socialness is negatively correlated with concreteness (Brysbaert et al., 2014), but also that the two variables share only a modest 10% of variance. Another key observation was that words rated as high in socialness spanned the entire concreteness dimension, from concrete concepts like *people* and *festival* to abstract ones like *democracy* and *cooperate*. As might be expected (Troche et al., 2014), socialness was positively associated with valence extremity measured as the absolute difference between the valence rating and the neutral point of the original valence scale (Warriner et al., 2013), but it shared only 4.8% of variance. We provide more extensive description and exploration of the socialness norms in Diveica et al. (2022) but, in summary, our preliminary analyses indicated that this socialness measure captures a distinct psycholinguistic construct.
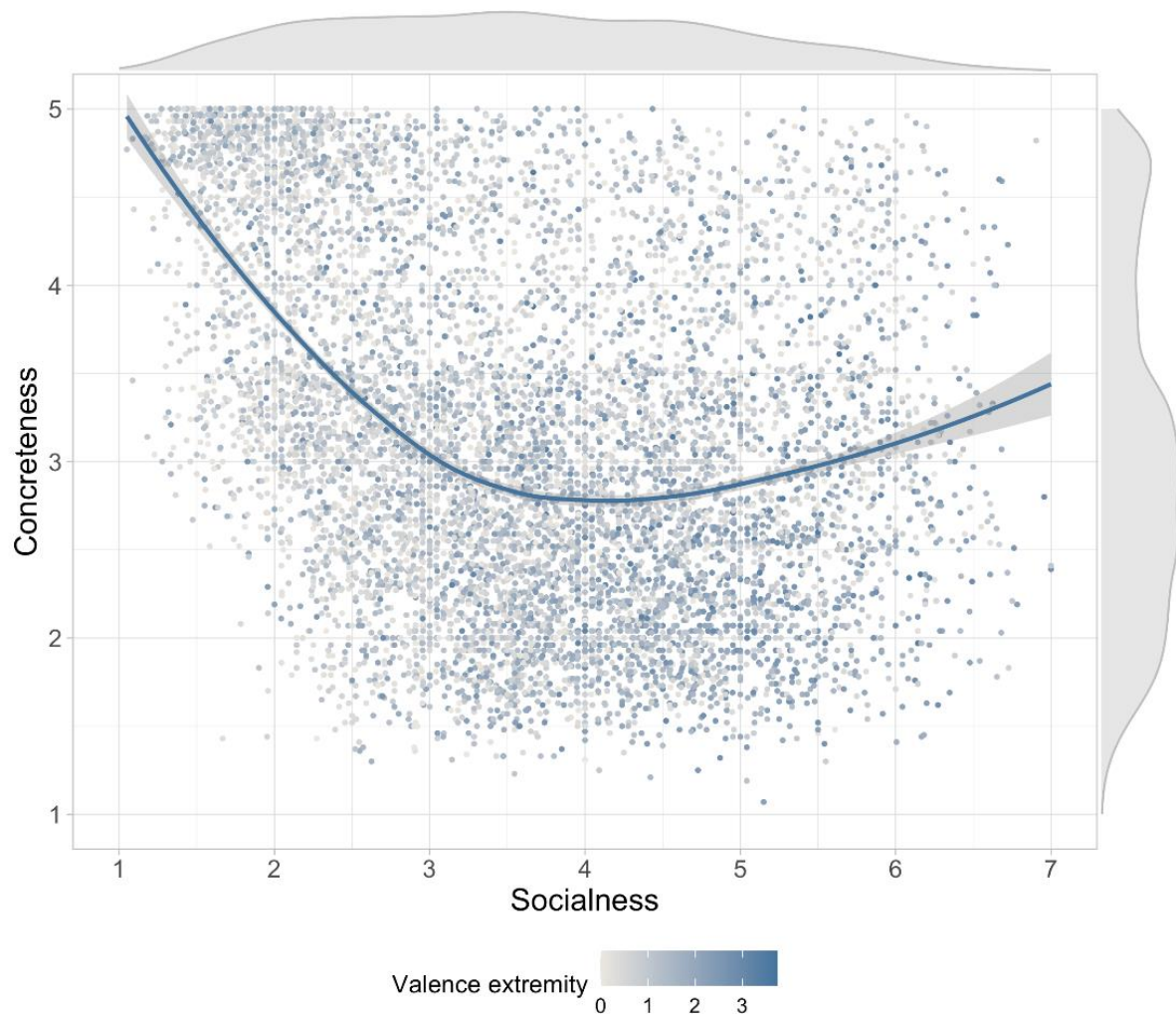
Figure 1. The relationship between socialness ratings (Diveica et al., 2022) and concreteness
ratings (Brysbaert et al., 2014) for 8,388 English words is illustrated and highlighted
by the loess line. The colour of the dots represents valence (Warriner et al., 2013)
extremity – the darker the colour, the more valenced the word. The density distributions
of the socialness and concreteness dimensions are plotted on the top and right of the
graph, respectively. The graph shows that words with high mean socialness ratings span
the entire concreteness dimension, and that the socialness measure captures information
distinct from valence.

5. Conclusions and future directions

The research we have reviewed here suggests that socialness, broadly construed, is a dimension
of word meaning that can be distinguished from other dimensions such as concreteness and
valence. Moreover, there is some evidence that socialness is reflected within the organisation
of neural systems that support semantic processing. It remains to be seen whether this is

indicative of social words being a distinct type, or whether socialness is just one of many dimensions that define a unified domain-general semantic space. At present, there remain two key shortcomings in this exciting area of research. First, researchers need to begin agreeing on terms and definitions of 'socialness' so that we are better able to compare theories and evaluate evidence in support of them (Gray, 2017). Second, there is very little research on the behavioural consequences of socialness and behavioural relevance is, of course, a gold standard for psychological theory. In terms of refining models of semantic representation, we believe that there are four key avenues for future research, and they have been made possible by the availability of the new socialness ratings (Diveica et al., 2022) described above. We will outline these research questions in the paragraphs below.

First, there are testable predictions that can be derived from WAT and other multiple representation theories. For instance, WAT proposes that social experience is key to learning and representing abstract concepts (Borghi et al., 2019). In line with this proposal, one could predict that (i) socialness facilitates the acquisition of abstract words, (ii) socialness contributes to the acquisition of abstract words more than to that of concrete words and (iii) abstract words are associated with more social content than are concrete words. In addition, WAT proposes a close link between linguistic and social experience. Consistent with this, Villani et al. (2019) found that, in a sample of abstract words, the more the words evoked social circumstances, the more they relied on auditory experiences, and on mouth motor system activation. These relationships could be further evaluated to understand how linguistic and social information jointly support acquisition and representation of abstract words.

Second, Diveica et al. (2022) characterized socialness in a broad and inclusive way and found this to be a useful and meaningful starting point. However, subsequent research is needed to more thoroughly explore the nature of the information captured by the socialness dimension and to evaluate whether there are important distinctions that it does not capture. In future research it will be helpful to consider narrower definitions to explore whether there are clusters or subtypes of social words. Moreover, it remains to be seen what aspects of the social experience, such as those measured by the more specific socially-relevant dimensions listed in Table 1, are most related to lexical-semantic processing, in terms of both behaviour and brain. It is possible that there are sub-types of social words that rely on different kinds of information. To some degree, this could mirror the more general concrete-abstract distinction, possibly in terms of how concepts rely differentially upon qualitatively different representational frameworks, such as those based on features/similarity and those based on association (cf., the proposal outlined by Crutch and Warrington in Crutch et al., 2007; Crutch & Warrington,

2005). For example, Roversi et al. (2013) investigated the properties associated with two potential sub-types of social concepts. They found that "social objects" (defined in Table 1), such as *choir*, elicited mainly contextual situations (e.g., *concert*), while institutional artefacts, such as *marriage*, evoked a higher proportion of normative relations (e.g., *commitment*). Further, the abstract-concrete distinction was more marked for social objects compared to institutional artefacts. Social objects that are concrete were associated with thematic/situational relations, while those that are abstract elicited more mental associations. In a related study, Villani et al. (2021) proposed a further distinction between pure institutional concepts (e.g., *marriage*) that relied more on exteroceptive information, and meta-institutional concepts (e.g., *duty*) for which interoceptive, affective and metacognitive information was more important. Future research that applies a data-driven approach across a large sample of abstract and concrete words will shed light on more specific socialness constructs and the way in which individual social word meanings potentially cluster together into sub-types.

Third, there are several implications for neuroimaging research into the representation of social concepts, and we have some recommendations. Now that large-scale socialness ratings are available and their independence from measures of concreteness and emotional valence has been more firmly established (Diveica et al., 2022), researchers are better positioned to comprehensively disentangle the neural correlates of socialness from other semantic variables. Indeed, right across the line of neuroimaging research reviewed in Part B, there is a need for greater integration of the kind of property generation, feature rating and behavioural research we reviewed in Part A. It will be instrumental for driving the next set of key questions, including those regarding the neural correlates of different types of concepts, and a putative privileged status afforded by socialness (Binney & Ramsey, 2020). At present, there is a lack of clear evidence in favour of an absolute boundary between social concepts and other types of concepts, and this suggests that there is going to be considerable overlap in the systems that represent them (Binney, Hoffman, et al., 2016; Rice et al., 2018). In this case, it will be important to use experimental designs and analytical techniques that allow for detecting more graded distinctions. To date, socialness has only been explored using univariate, magnitude-based approaches, whereas information-based approaches, including multivariate pattern analysis and repetition suppression paradigms, will be essential, particularly for understanding overlapping activation, which could reflect either shared processes, or tightly yet separately packed cognitive functions that only dissociate when investigated at higher spatial resolutions (Henson, 2006; Humphreys et al., 2020). Moreover, a key methodological determinant for obtaining a complete picture of the neural basis of social concepts will be the

use of neuroimaging techniques that maximise the signal obtained from across the entirety of key brain regions. This includes the anterior temporal lobe, of which some subregions are invisible to standard fMRI (Balgova et al., 2021; Binney, Hoffman, et al., 2016; Embleton et al., 2010).

Fourth, it is worth noting that concepts are not static and that their representation depends on ongoing task contexts as well as prior experience (Pexman, 2020; Yee & Thompson-Schill, 2016). For example, it has been shown that concepts are to some degree influenced by culture and the language spoken (Malt & Majid, 2013; Thompson et al., 2020). Given that cultural environments are intrinsically linked to our social experiences, social concepts might be particularly susceptible to cultural influences. Moreover, a variety of socially relevant characteristics (e.g., race, gender) impact our social experiences, which can consequently lead to between-individual variability in the representation of social concepts. In line with this, Mazzuca et al. (2020) showed that the features participants associated most strongly with the social concept *gender* depended on their gender identity and sexual orientation. This potential variability could be investigated in future research and might manifest in various ways. For instance, some abstract words, including those high in social content, might place greater demands upon cognitive control processes because their exact meaning is dependent on context. This might be reflected in differential engagement of regions associated with controlled semantic selection and retrieval, such as the left inferior frontal gyrus (IFG; for related discussions, see Diveica et al., 2021; Hoffman et al., 2015; Hoffman & Tamm, 2020). Consistent with this, some individual neuroimaging studies reported greater activation of semantic control regions (the IFG) during the processing of social, as compared to non-social words/sentences (Binney, Hoffman, et al., 2016; Mellem et al., 2016; also see Satpute et al., 2014). In addition, there is some limited behavioural evidence that implicit semantic processing of social words compared to non-social words slows reaction times in a Stroop task in adults (Arioli, Basso, et al., 2021) and in a selective attention task in children (Pérez-Edgar & Fox, 2007), indicating a greater demand for cognitive control. However, more research is needed to understand what task contexts and concept features might drive an increased need for regulatory mechanisms when processing social concepts.

In summary, in the present paper we have outlined the ways in which two different literatures have explored the idea that social concepts might be a special type and have offered suggestions for integrating and advancing these research efforts. Further, we presented initial psycholinguistic explorations of a new and openly available set of socialness ratings for over 8,000 words (for a detailed description, see Diveica et al. 2022). These suggest that socialness

is indeed a distinct aspect of word meaning, and one that should be incorporated in theories of semantic representation. Social words, like *manner*, *gentleman*, and *refuse*, convey information about our relationships with people, and inform our understanding of their actions. Socialness gives words salience and gives meaning to the interactions and events that make up sources like *Pride and Prejudice*, and that occur in the personal and interpersonal complexities of our everyday lives.

CHAPTER 3

Quantifying social semantics: an inclusive definition of socialness and ratings for 8,388 English words

This chapter has been published as:

Authors' contributions:

Veronica Diveica, Dr Richard Binney and Prof Penny Pexman conceptualized this study, developed the methodological approach and interpreted the results. Veronica Diveica performed the data collection, data curation, data visualization and statistical analyses, and wrote the original draft. Dr Richard Binney and Prof Penny Pexman edited drafts of the paper.

Open science practices and data/code availability statement:

The datasets generated and/or analyzed in the current study and the analysis scripts are available via the Open Science Framework (OSF) project: https://osf.io/2dqnj/. Given its exploratory nature, the study was not pre-registered.

Abstract

It has been proposed that social experience plays an important role in the grounding of concepts, and *socialness* has been proffered as a fundamental organisational principle underpinning semantic representation in the human brain. However, the empirical support for these hypotheses is limited by inconsistencies in the way socialness has been defined and measured. To further advance theory, the field must establish a clearer working definition, and research efforts could be facilitated by the availability of an extensive set of socialness ratings for individual concepts. Therefore, in the current work we employed a novel and inclusive definition to test the extent to which socialness is reliably perceived as a broad construct, and we report socialness norms for over 8000 English words, including nouns, verbs and adjectives. Our inclusive socialness measure shows good reliability and validity, and our analyses suggest that the socialness ratings capture aspects of word meaning which are distinct to those measured by other pertinent semantic constructs, including concreteness and emotional valence. Finally, in a series of regression analyses, we show for the first time that the socialness of a word's meaning explains unique variance in participant performance on lexical tasks. Our dataset of socialness norms has considerable item overlap with those used in both other lexical/semantic norms and in available behavioural mega-studies. They can help target testable predictions about brain and behaviour derived from multiple representation theories and neurobiological accounts of social semantics.

1. Introduction

Conceptual knowledge is the foundation of our complex interactions with the environment, bringing meaning to the objects, words and social agents we encounter. A major challenge for the cognitive sciences is therefore to characterise how meaning is represented in the brain. Of particular interest has been the issue of how the mental representations of concepts become connected to their referents, termed the symbol grounding problem (Harnad, 1990; Searle, 1980). Within multiple representation accounts of semantic processing, concepts are mapped to the world, or *grounded*, by being directly represented within the neural systems underpinning multiple experiential channels such as perception, action, emotion, language and cognition (Borghi et al., 2018; Kiefer & Harpaintner, 2020). Sensorimotor systems are particularly important for grounding concrete concepts such as *festival* and *politician.* In contrast, abstract concepts like *romance* and *democracy* cannot, by definition, be directly experienced through the senses, and may thus rely to a greater degree on other types of information, such as affective (Fingerhut & Prinz, 2018; Kousta et al., 2011), introspective (Shea, 2018) and linguistic experience (Borghi et al., 2019; Dove, 2018). Further, there is growing recognition that there are different types of abstract concepts which depend to varying extents on these manifold sources of information (Harpaintner et al., 2018; Villani et al., 2019) and which elicit different patterns of behavioural responses in lexical-semantic tasks (Muraki et al., 2020).

Recently, there has been a rise in interest concerning the role that social experience plays in the acquisition and representation of concepts. Indeed, there are proposals in which social interaction and social context are pinpointed as a key source or mechanism for grounding that may be particularly important for the representation of abstract concepts (Barsalou, 2020; Borghi et al., 2019). For instance, Barsalou (2020) proposed that the social environment (e.g., agents, social interaction, culture) provides one form of grounding, in addition to that afforded by perceptual modalities, both of which are distinguished from the body, and the physical environment. Likewise, Borghi et al. (2019) argued that both social interactions and linguistic inputs are crucial for the acquisition of abstract concepts (also see Borghi & Binkofski, 2014). In Pexman et al. (2021), we have reviewed these theoretical perspectives as well as two parallel sets of empirical literature which provide some evidence for socialness being a key principle underpinning semantic representation. For example, property generation and feature ratings studies found that social semantic content, or *socialness*, helps distinguish concrete from abstract concepts (Barsalou & Wiemer-Hastings, 2005; Troche et al., 2014; Wiemer-Hastings

& Xu, 2005) and even different sub-types of abstract concepts (Harpaintner et al., 2018; Villani et al., 2019). In parallel, a set of neuroimaging studies have found that words high in socialness are associated with differential patterns of brain activation during semantic processing (e.g., Arioli et al., 2021; Binney et al., 2016; Mellem et al., 2016; Rice et al., 2018; Wang et al., 2019; for another review, also see Conca et al., 2021). Some authors have argued for a special status of social concepts over other types of concepts, and have suggested that socialness may even be a fundamental driver behind the functional organisation of the semantic system (Lin et al., 2018; Ross & Olson, 2010; Simmons et al., 2010; Zahn et al., 2007). These studies were all based on limited word samples, but they provide some evidence that social words might be a distinct type of concept, in line with proposals of some multimodal (e.g., Borghi et al., 2018; Kiefer & Harpaintner, 2020) and neurobiological models (e.g., Olson et al., 2013) of conceptual processing.

These theories are nascent and there are many outstanding questions about the nature and extent of the contribution that socialness makes to semantic representation. One fundamental question is whether socialness is a behaviourally relevant principle as indexed, for example, by its ability to account for variance in performance on lexical-semantic tasks. However, the extant empirical support is limited by the way socialness has been defined and measured. To our knowledge, the largest source of openly available socialness norms was compiled by Troche et al. (2017) and includes social interaction ratings for 750 English nouns. Another dataset collected by Binder et al. (2016) includes ratings for 434 nouns, 62 verbs, and 39 adjectives on four socially-relevant dimensions labelled social, communication, human and self. Thus, the scale and scope (i.e., the syntactic classes of words) at which socialness has been explored has been limited to date. Moreover, *socialness* as a construct has been defined variably in terms of behavioural descriptiveness, and there is no consensus on the criteria that differentiate social from non-social concepts. The heterogeneity in definitions is summarised by Pexman et al. (2021); some researchers have measured socialness as, for example, the degree to which a word's meaning refers to relationships between people (Troche et al., 2014, 2017), to social as opposed to individual contexts (Arioli, Basso, et al., 2021), or to the relationship between self and others (Crutch et al., 2012), and socialness has also been defined as how well words describe social behaviour (Zahn et al., 2007). This variability in the operationalisation of socialness hinders our ability to compare findings across studies and glean a broader understanding of the contribution made by socialness to conceptual representation in the brain, and its behavioural consequences. Thus, we argue that to further progress theory, the field must first establish a clearer working definition of socialness.

Moreover, many of these past studies employed socialness definitions that emphasise specific aspects of social experience (Pexman et al., 2021). These narrow definitions might neglect important aspects of our highly complex interactions with the social environment. Thus, taking a crucial next step for understanding the construct of *socialness,* we aimed to collect ratings using an inclusive definition designed to capture all manner of features that are deemed to be socially relevant. This allowed us to test the extent to which socialness is reliably perceived as a broad construct. Relatedly, our socialness definition can be equally applied to a wide range of words, from nouns like those referring to social roles (e.g., *lawyer*) or institutions (e.g., *government*), to verbs like *to befriend*, and adjectives like *trustworthy*. This broad and inclusive definition can be used as a starting point for future studies exploring more fine-grained aspects of the socialness construct.

In summary, the aims of the present study were as follows: 1) collect socialness ratings for a large set of English words to provide a useful resource for future research endeavours; 2) use an inclusive definition to assess the extent to which socialness is reliably perceived as a broad construct; 3) explore to what extent these new socialness ratings capture aspects of word meaning that are distinct from those measured via other related semantic variables, such as concreteness and emotional valence, and 4) test whether socialness is a behaviourally relevant construct.

2. Methods

2.1. Participants

Participants were recruited via the online platform Prolific (https://www.prolific.co/). Responders were restricted to those who self-reported being fluent in English and having no language disorders. A total of 605 participants (359 male, 240 female, six unspecified, M_{age} = 29.44 years, SD_{age} = 10.6) completed the study. Participants completed the rating task in 34 minutes on average and were compensated with GBP £4. Following exclusions (see below), the final sample consisted of 539 participants, with ages ranging from 18 to 76 years (M = 29.7; SD = 10.67). Of the participants, 216 (40.07%) were female, 317 (58.81%) male and six (1.11%) unspecified. English was the first language for 273 (50.65%) participants. Of the remaining 266 (49.35%) participants, 111 self-reported as being proficient in English, 124 advanced and 31 beginner/intermediate. A total of 185 (34.32%) participants were monolingual, while the remaining 354 (65.68%) reported speaking more than one language.

2.2. Stimuli

The stimuli were 8948 words, including 5569 nouns, 1343 verbs, 2009 adjectives, and 26 other parts of speech (based on the dominant part-of-speech norms in Brysbaert et al., 2012) [1]. We compiled our stimulus set from two sources: the Calgary Semantic Decision Project (Pexman et al., 2017) and Brysbaert et al. (2014)'s dataset of concreteness ratings. Ratings on emotion dimensions (valence, arousal, dominance) from Warriner et al. (2013) and on concreteness from Brysbaert et al. (2014) are available for all of the words included and the selected words span the entire continuum of these dimensions. In addition, we specifically selected these words so that there would be considerable overlap with behavioural mega-studies and other theoretically important psycholinguistic dimensions, some of which were used in analyses reported below, whereas others might be of interest in future research (e.g., Calgary Semantic Decision Project (Pexman et al., 2017), the Lancaster Sensorimotor Norms (Lynott et al., 2020), the Glasgow norms (Scott et al., 2019), word association norms (De Deyne et al., 2019), word prevalence norms (Brysbaert et al., 2018)).

We used 30 of the 8,948 words as a set of control items which were to be presented to every participant and used during the data cleaning process (see below). These words were selected based on the ratings received in a pilot study ($N = 36$ participants) that was run to obtain an initial assessment of whether participants understand the task instructions and, in particular, the description of the inclusive socialness measure, and whether they provide reliable ratings (for a detailed description, see Section S1 of Supplementary Materials). Control words were selected to vary in the mean pilot socialness ratings, as well as in their concreteness (Brysbaert et al., 2014) and valence ratings (Warriner et al., 2013).

In addition to the 8948 words, we selected 12 practice words to be rated before the main ratings task so that participants could become familiar with the task requirements. We selected practice words that vary in concreteness (Brysbaert et al., 2014) and valence (Warriner et al., 2013), and that span the whole range of the social interaction dimension as measured by Troche et al., (2017) to ensure that participants practised both items with high and with low socialness ratings.

We used Qualtrics software (Qualtrics, 2020) to create two questionnaires for presentation to participants. To facilitate efficient Qualtrics processing, we divided the 8918 words into two lists of 4459 words from which each participant saw a random subset. These

---

[1]Note that part-of-speech information was not available for one word: *hip hop*.

lists were equated for letter length, frequency (log subtitle frequency; Brysbaert & New, 2009), concreteness (Brysbaert et al., 2014) and valence (Warriner et al., 2013) to control for the probability of selecting words with different characteristics from each list. The control words were then added to both lists, resulting in two questionnaires each with 4489 words.

2.3.   Procedure

The word stimuli were presented using Qualtrics (2020) and linked to the Prolific online recruitment platform (www.prolific.co). Following the consent form, a demographics survey and instructions, participants rated the 12 practice words, then proceeded to rate the main set of items. Each participant rated 370 words randomly selected from one of the two item lists, plus the 30 control words. The control words were randomly intermixed with other items. The full instructions given to participants are presented in Section S2 of supplementary materials. In short, the participants were asked to rate the degree to which the words' meaning has social relevance by describing or referring to the following:

> *a social characteristic of a person or group of people, a social behaviour or interaction, a social role, a social space, a social institution or system, a social value or ideology, or any other socially-relevant concept.*

Participants provided their answers using a seven-point Likert scale presented horizontally below each word. In addition, there was an "I don't know the meaning of this word" option. There were 25 words presented per page. We collected data until we obtained at least 25 ratings per word.

2.4.   Data cleaning

In total, we collected 241,575 observations. The data cleaning pipeline involved sequentially implementing several techniques consistent with recommendations for identifying careless or insufficient effort responders (Curran, 2016) and computer-generated random responding (Dupuis et al., 2019), as well as other data cleaning procedures used in previous word norming studies (Brysbaert et al., 2014; Pexman et al., 2019; Warriner et al., 2013). First, we removed data from participants if they completed less than 33% of the ratings task ($n = 0$), responded with "I don't know the meaning of this word" for more than 25% of items ($n = 8$) and provided the same rating for more than 25 words in a row ($n = 17$). Next, we examined each participant's ratings of the 30 control words and generated correlations with the mean ratings of those words obtained in the pilot study. We removed data from 36 participants with a correlation coefficient less than .20. We then computed the correlation between each participant's ratings and the

mean ratings of all other participants. We deleted data from five participants with a correlation coefficient less than .10. Finally, if more than 15% of raters reported not knowing a particular word, we removed those words from the analyses reported below. This led to the exclusion of 560 words.

The final dataset was comprised of 8388 words and 202,841 observations, of which 3542 were "I don't know the meaning of this word" responses. Not taking into account the control words rated by all participants, each word in the final dataset had 21.92 valid ratings on average ($SD = 1.68$), ranging from 15 to 27 ratings. Overall, 7703 (91.83%) words had at least 20 valid ratings.

2.5.   Data analysis overview

Data pre-processing, analysis and visualisation was accomplished using R version 3.6.1 (RStudio Team, 2020). We first computed descriptive statistics for the socialness ratings and assessed their reliability. Then, to begin to explore the nature of the information captured by the socialness dimension and characterize its relationship with other pertinent psycholinguistic constructs, we computed the zero-order correlations between the mean socialness ratings and a variety of lexical and semantic properties of the words. Next, we conducted a series of hierarchical regression analyses to examine whether the socialness measure is related to behaviour in lexical tasks, using behavioural responses from the English Lexicon Project (ELP) lexical decision task (Balota et al., 2007) and the English Crowdsourcing Project (ECP) word knowledge task (Mandera et al., 2020). The LDT outcome variables quantify the speed and accuracy with which participants could distinguish between words and non-word letter strings. The ECP RT outcome variable measures the speed with which participants could recognize a word as known to them, while the percentage of participants reporting not knowing a word (henceforth proportion unknown) is a measure of word prevalence. We selected these tasks because they require only a fairly shallow level of semantic access (Muraki et al., 2020) and thus provide a conservative test of the relationship between this measure and lexical semantic processing. In addition, in both of these tasks, all word stimuli received the same behavioural response ("word" in the ELP LDT, or "I know that word" in the ECP) unlike, for instance, semantic decision tasks (e.g., Pexman et al., 2017) which involve different responses for different types of words. All predictor variables were mean-centred and we used reaction times standardized as z-scores because these reduce the influence of individual differences on overall processing speed (Faust et al., 1999).

3. Results

3.1. Descriptive statistics

The raw data and resulting socialness ratings are provided on the Open Science Framework (OSF) project page (available at: https://osf.io/2dqnj/). The socialness ratings have a unimodal distribution with a mean of 3.63 ($SD = 1.24$) (Figure 1a). More descriptive statistics for the mean ratings are provided in Table 1 and the distribution of ratings as a function of part of speech is depicted in Figure 1b. The ratings have an average standard deviation of 1.85 ($SD = 0.35$) and participants provided more consistent responses at the extremes of the scale (Figure 1c). Examples of words at the extremes of the socialness dimension are given in Table 2. Words like *friendship*, *people* and *sociable* received high socialness values, while words like *avalanche*, *millimeter* and *hemoglobin* received low socialness ratings, suggesting good face validity.

Table 1. Descriptive Statistics for Socialness Ratings for 8388 Words.

Descriptive Statistic	Value
Mean	3.63
Median	3.57
Standard Deviation	1.24
Minimum	1.05
Maximum	7.00
1st Quartile	2.62
3rd Quartile	4.58
Skewness	0.19
Kurtosis	-0.80

Table 2. List of words at the extremes of the socialness dimension.

Highest-Rated Words	Rating	Lowest-Rated Words	Rating
friendship	7.00	eucalyptus	1.05
socialize	7.00	horizontal	1.09
relationship	6.96	crocodile	1.09
people	6.90	sulfur	1.10
romance	6.78	sleeve	1.17
marriage	6.76	turbo	1.18
socialism	6.75	cranberry	1.18
political	6.73	dragonfly	1.18
family	6.72	hemoglobin	1.20
teamwork	6.72	shark	1.21
boyfriend	6.68	sunflower	1.21
friend	6.68	sandpaper	1.22
sociable	6.68	millimeter	1.22
sisterhood	6.67	avalanche	1.22
mother	6.67	spinach	1.22
democracy	6.65	airspeed	1.23
togetherness	6.65	button	1.23
sister	6.65	redwood	1.23
festival	6.64	pistachio	1.24
stepfather	6.64	birch	1.25
humankind	6.62	haystack	1.25
meeting	6.62	toothpaste	1.26
parental	6.62	paprika	1.27
befriend	6.61	cellophane	1.28
chatty	6.61	magnolia	1.28
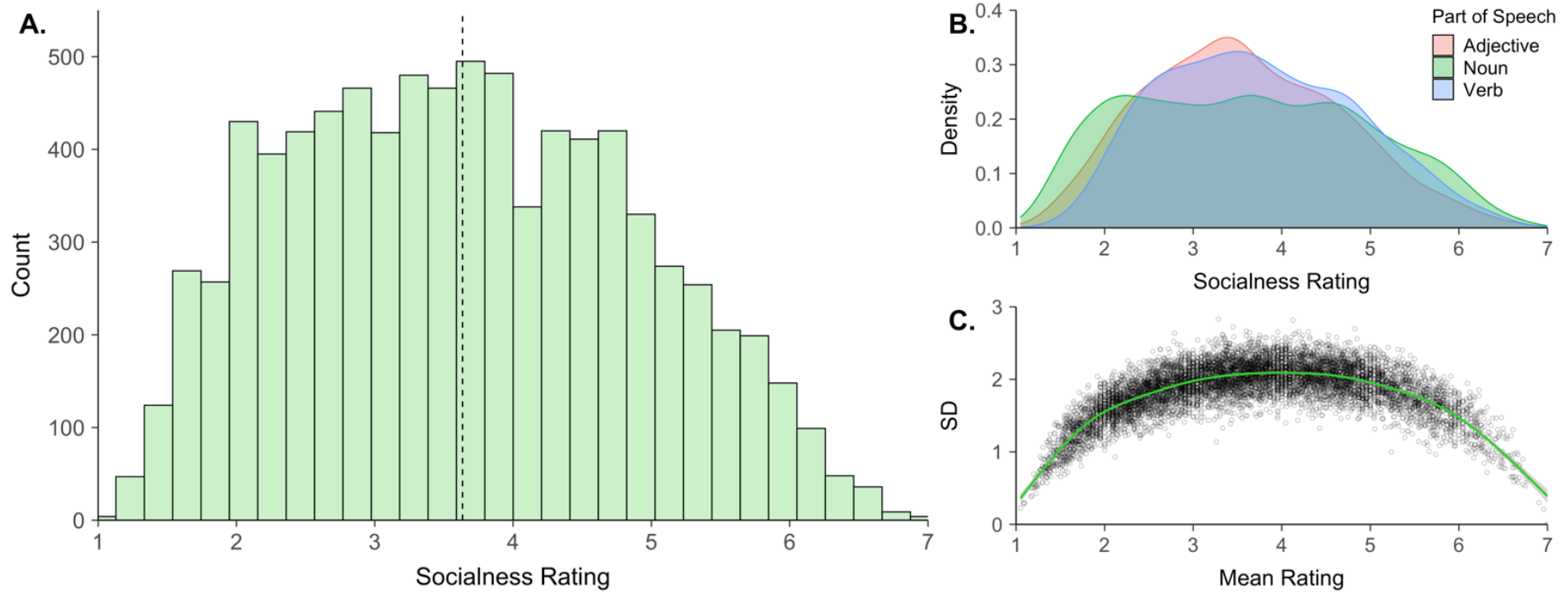
Figure 1. Distribution of socialness ratings. A. Histogram of socialness ratings for 8388 words; the dotted line represents the mean. B. Kernel density plot of ratings as a function of syntactic class. C. Standard deviation of ratings plotted against their respective mean rating, along with a loess line (in green) that highlights the functional relationship.

3.2. Reliability and validity

We first examined the reliability of the ratings by computing the one-way intra-class correlation coefficient (ICC) of all ratings using variances estimated via a random effects model with a global intercept and a random intercept per word (Brysbaert, 2019; Stevens & Brysbaert, 2016). We found an ICC of 0.9 which indicates good reliability of the mean socialness ratings. We further computed the split-half reliability for the 30 control words which were the only items in our dataset rated by all participants. We found a mean Spearman–Brown corrected split-half reliability of 0.998 ($SD = 0.16$) across 100 random splits, suggesting high reliability for the control items.

We then examined the validity of the ratings by computing the correlations between the ratings observed here and the mean ratings collected in the pilot study ($n = 60$ words), as well as two previous related sets of social interaction norms collected by Binder et al. (2016) ($n = 258$ words), and Troche et al. (2017) ($n = 450$ words). The current socialness ratings were strongly and positively correlated with the ratings collected in the pilot study ($r = 0.97$) and with the previous social interaction ratings collected by Binder et al. (2016) ($r = 0.76$) and Troche et al. (2017) ($r = 0.76$), suggesting good validity.

3.3.   Correlations with lexical and semantic properties

We examined the correlations between the socialness ratings and various lexical and semantic properties of the words. We included lexical dimensions in our analysis as previous work has shown that semantic content is not independent of the linguistic properties of words (Lewis & Frank, 2016; Reilly et al., 2012, 2017; Strik-Lievers et al., 2021). The lexical variables included letter length, orthographic Levenshtein distance (Yarkoni et al., 2008), phonological Levenshtein distance and frequency (log subtitle frequency; Brysbaert & New, 2009). To examine the proposed relationship between socialness and abstractness (Borghi et al., 2019), we included the following semantic variables that index sensorimotor experience: concreteness (Brysbaert et al., 2014), imageability (Cortese & Fugett, 2004; Schock et al., 2012), body-object interaction (BOI; the ease with which a human body can physically interact with a word's referent; Pexman et al., 2019), and sensory experience ratings (Juhasz & Yap, 2012). To assess the generalizability of the association between socialness and affective information reported in previous studies (Troche et al., 2014, 2017; Villani et al., 2019), we included in our analysis valence extremity (the degree to which the word evokes positive/negative feelings; this was measured as the absolute difference between the valence rating and the neutral point

of the original valence scale by Warriner et al., 2013), arousal (Warriner et al., 2013), and dominance (Warriner et al., 2013). Finally, to assess the relationship between the socialness ratings and linguistic experience, the semantic variables included semantic diversity (Hoffman et al., 2013), rating-based age of acquisition (AoA) (Kuperman et al., 2012), and a test-based AoA measure derived from Dale and O'Rourke (1981) and updated by Brysbaert and Biemiller (2017).

These correlations revealed several interesting relationships that provide insight as to the nature of the word socialness measure (Figure 2; see Figure S1 for scatterplots). Socialness was negatively correlated with concreteness ($r = -0.32$), imageability ($r = -0.18$), and BOI ($r = -0.17$), which suggests that words with less social relevance are associated with more embodied sensorimotor information. In contrast, socialness ratings were positively correlated with valence extremity ($r = 0.22$) and arousal ($r = 0.22$), suggesting that social words tend to have more affective information.

	Length	Frequency	OLD	PLD	AoA Rating	AoA Test-based	Concreteness	Imageability	BOI	SER	Valence Extremity	Arousal	Dominance	Semantic Diversity
Dominance														0.11
Arousal													-0.2	
Valence Extremity												0.36	-0.18	
SER											0.24	0.24	-0.09	-0.28
BOI										0.36	-0.08	-0.17	0.11	-0.39
Imageability									0.68	0.62	0.07			-0.34
Concreteness								0.76	0.79	0.41	-0.15	-0.14		-0.39
AoA Test-based							-0.32	-0.36	-0.15	-0.18	-0.11	0.04	-0.14	-0.12
AoA Rating						0.72	-0.33	-0.41	-0.16	-0.25	-0.13	0.05	-0.17	-0.16
PLD					0.29	0.19	0.03	0.14	0.33	0.18		0.05	-0.05	-0.19
OLD				0.89	0.27	0.17	0.06	0.18	0.35	0.23		0.05	-0.07	-0.2
Frequency			-0.37	-0.37	-0.53	-0.48		0.09	-0.17		0.14	0.04	0.15	0.34
Length		-0.35	0.85	0.81	0.28	0.16		0.2	0.32	0.21		0.04		-0.13
Socialness	0.12	0.18	0.06	0.07	0.1	0.04	-0.32	-0.18	-0.17	-0.09	0.22	0.22	0.07	0.08

Pearson r: 1.0, 0.5, 0.0, -0.5, -1.0

Figure 2. Correlations between mean socialness ratings and lexical-semantic dimensions. Only correlations significant at *p* < .01 are shown. The strength and direction of the correlation coefficients are indicated by the colour and the numerical values. For each variable of interest, the numbers of items in common with our socialness ratings are as follows: length, concreteness, valence, arousal and dominance: 8,388; log subtitle frequency: 8160; OLD and PLD: 8027; rating-based AoA: 8348; test-based AoA: 7321; imageability: 2680; BOI: 4038; SER: 2645. SER = sensory experience rating; BOI = body-object interaction; AoA = age of acquisition; PLD = phonologic Levenshtein distance; OLD = orthographic Levenshtein distance.

3.4. Relationships with performance on lexical tasks

Next, we examined whether the socialness ratings are related to lexical-semantic processing using behavioural responses from the ELP LDT (Balota et al., 2007) and the ECP word knowledge task (Mandera et al., 2020). We conducted a series of item-wise hierarchical regression analyses in which we included other lexical and semantic predictors (that are

typically related to behaviour in lexical tasks) in order to isolate the unique relationships of socialness to standardized reaction times (RTs), ELP error rates and ECP proportion unknown. In the first step, we entered the control predictors letter length, frequency (Brysbaert & New, 2009) and rating-based AoA (Kuperman et al., 2012). In the second step, we entered the semantic predictors: socialness, concreteness (Brysbaert et al., 2014), valence extremity (Warriner et al., 2013) and semantic diversity (Hoffman et al., 2013). We selected these other semantic predictors on the basis of multidimensional theories (e.g., Borghi et al., 2019) that highlight the simultaneous contribution of semantic variables derived from multiple sources, including linguistic (semantic diversity), sensorimotor (concreteness) and affective experience (valence extremity).

There were 6926 items for which we had values for all variables of interest in the analysis predicting LDT performance. Descriptive statistics and zero-order correlations between all variables of interest from this dataset are reported in Supplementary Table S1. The statistical results are reported in Table 3 and the standardized coefficients are illustrated in Figure 3a. In this analysis, the control variables were all significant predictors of LDT latencies – RTs were faster for words that are shorter, more frequent and acquired earlier. There was significant improvement in model fit with the addition of the semantic variables, which collectively accounted for a further 0.61% of variance in LDT latencies. Of the semantic variables, only socialness and semantic diversity were significant predictors, with faster RTs for words with increased social relevance and for those encountered in more semantically diverse contexts. A similar pattern of results was observed when predicting LDT error rates. The control variables were all significant predictors, with fewer errors for words that are longer, more frequent and acquired earlier. There was significant improvement in model fit with the inclusion of the semantic variables, which accounted for an additional 0.56% of variance in LDT error rates. Socialness and semantic diversity were the only significant semantic predictors – error rates were lower for words with increased socialness and for those that are more semantically diverse.

Table 3. Regression coefficients from item-level analyses predicting lexical decision task latencies and error rates (N = 6926).

Predictor	zRTs							Error Rates						
	b	SE	t	p	sr^2	R^2	ΔR^2	b	SE	t	p	sr^2	R^2	ΔR^2
Step1						0.51							0.21	
Intercept	-0.25	0.003	-94.49	***				0.06	0.001	70.97	***			
Length	0.05	0.001	35.6	***	0.09			-0.01	< .001	-22.57	***	0.058		
Frequency	-0.15	0.005	-29.99	***	0.064			-0.03	0.002	-19	***	0.041		
Age of Acquisition	0.04	0.001	26.91	***	0.051			0.01	< .001	22.99	***	0.06		
Step2						0.52	0.006						0.22	0.006
Intercept	-0.25	0.003	-95.06	***				0.06	0.001	71.21	***			
Length	0.05	0.001	35.75	***	0.089			-0.01	< .001	-21.5	***	0.052		
Frequency	-0.13	0.005	-23.9	***	0.04			-0.03	0.002	-14.7	***	0.024		
Age of Acquisition	0.04	0.001	25.78	***	0.046			0.01	0.001	22.31	***	0.056		
Socialness	-0.01	0.002	-4.73	***	0.002			-0.003	0.001	-3.57	***	0.001		
Concreteness	< .001	0.004	0.02	0.984	0			0.002	0.001	1.7	0.088	0		
Valence Extremity	0.01	0.004	1.83	0.067	0			-0.001	0.001	-0.64	0.525	0		
Semantic Diversity	-0.07	0.01	-6.77	***	0.003			-0.01	0.003	-3.54	***	0.001		

Note. b represents unstandardized regression weights. SE represents the standard error of the regression weights. sr^2 represents the semi-partial correlation squared. LDT lexical decision task. zRTs standardized reaction times. *p < .05; **p < .01; ***p < .00

There were 7010 items for which we had values for all variables of interest in the analysis predicting performance in the ECP word knowledge task. Descriptive statistics and zero-order correlations between all variables of interest from this dataset are reported in Supplementary Table S2. The statistical results are reported in Table 4 and the standardized coefficients and illustrated in Figure 3b. In this analysis, the control variables were all significant predictors of response latencies – RTs were faster for words that are shorter, more frequent and acquired earlier. There was significant improvement in model fit with the addition of the semantic variables, which accounted for a further 0.78% of variance in recognition RTs. All semantic variables were significant predictors, with faster RTs for words with increased socialness, concreteness and valence extremity and for those encountered in more semantically diverse contexts. The control variables were all significant predictors of the proportion of people reporting not knowing a word, with words that are longer, more frequent and acquired earlier being more prevalent. There was significant improvement in model fit with the inclusion of the semantic variables, which accounted for an additional 0.83% of variance in ECP proportion unknown. Valence and semantic diversity were the only significant semantic predictors – words that are more valenced and encountered in more semantically diverse contexts were reported as known by more people.

Table 4. Regression coefficients from item-level analyses predicting ECP word knowledge task latencies and proportion unknown (N = 7010).

Predictor	zRTs							Proportion Unknown						
	b	SE	t	p	sr^2	R^2	ΔR^2	b	SE	t	p	sr^2	R^2	ΔR^2
Step1						0.4							0.23	
Intercept	-0.53	0.001	-495.33	***				0.013	< .001	69.57	***			
Length	0.01	0.001	19.37	***	0.032			-0.002	< .001	-22.24	***	0.055		
Frequency	-0.06	0.002	-27.65	***	0.065			-0.007	< .001	-19.99	***	0.044		
Age of Acquisition	0.01	0.001	25.28	***	0.054			0.002	< .001	24.02	***	0.064		
Step2						0.41	0.008						0.23	0.008
Intercept	-0.53	0.001	-498.44	***				0.013	< .001	69.93	***			
Length	0.01	0.001	20.21	***	0.034			-0.002	< .001	-21.67	***	0.051		
Frequency	-0.05	0.002	-22.07	***	0.041			-0.006	< .001	-15.74	***	0.027		
Age of Acquisition	0.01	0.001	22.5	***	0.043			0.002	< .001	22.38	***	0.055		
Socialness	-0.003	0.001	-3.6	***	0.001			< .001	< .001	-0.31	0.754	0		
Concreteness	-0.003	0.001	-2.04	*	< .001			< .001	< .001	1.46	0.145	0		
Valence Extremity	-0.01	0.001	-6.09	***	0.003			-0.001	< .001	-3.52	***	0.001		
Semantic Diversity	-0.02	0.004	-6.01	***	0.003			-0.004	0.001	-5.89	***	0.004		

Note. b represents unstandardized regression weights. SE represents the standard error of the regression weights. sr^2 represents the semi-partial correlation squared. zRTs standardized reaction times. *p < .05; **p < .01; ***p < .001
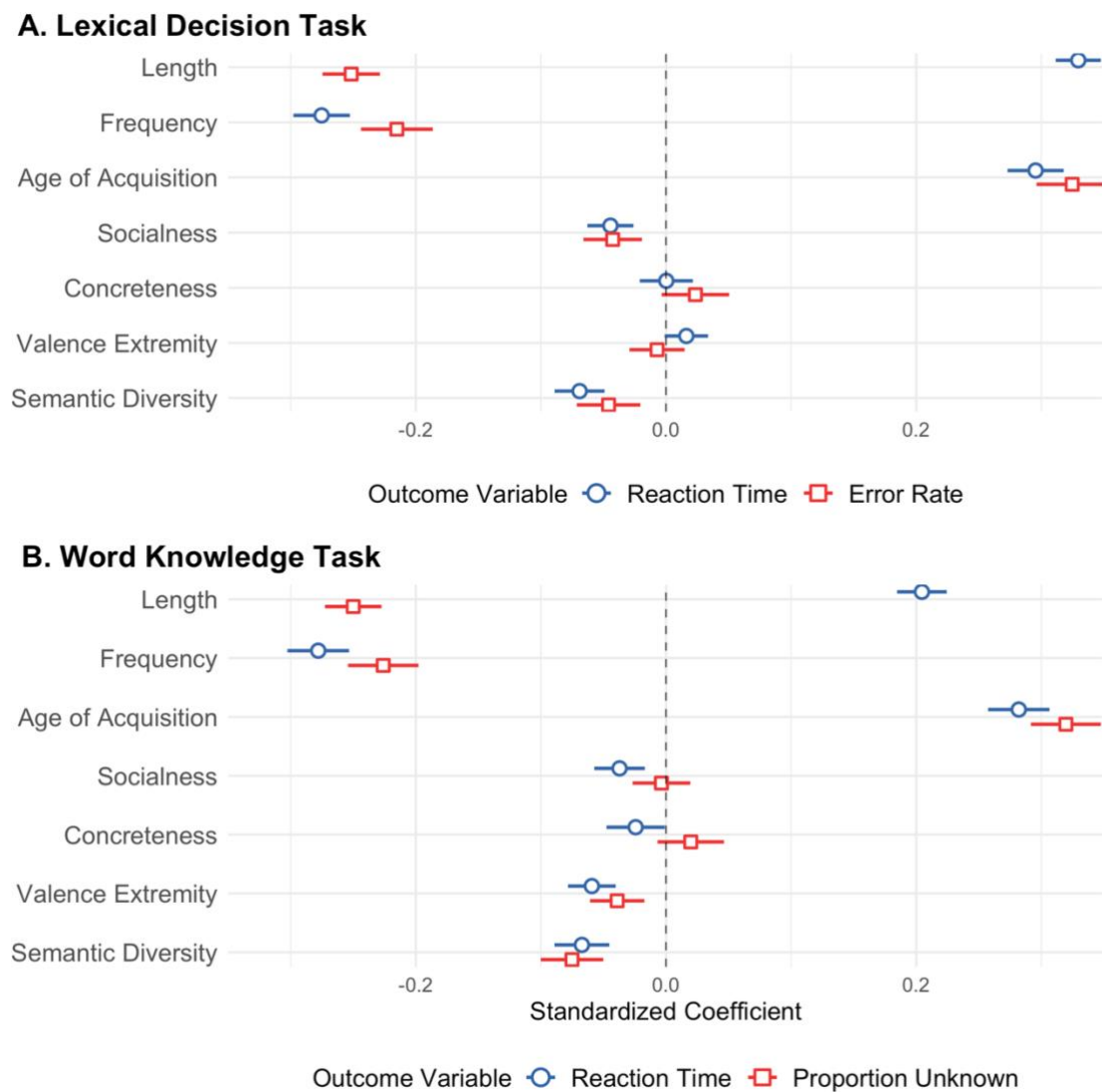
Figure 3. Standardized coefficient weights and 95% CIs for the second step of the hierarchical regression analyses predicting task outcome variables. A. Standardized beta coefficients for LDT RTs (blue) and errors (red). B. Standardized beta coefficients for ECP Word Knowledge Task RTs (blue) and the proportion of people reporting not knowing a word (red).

4. Discussion

Although some contemporary accounts (e.g., Barsalou, 2020; Borghi et al., 2019; Kiefer & Harpaintner, 2020) proffer a role for socialness in the organization and grounding of conceptual knowledge, many key questions remain about the nature of its contribution and its neural underpinnings. With the aim of facilitating future endeavours, in the present work we sought

to 1) collect socialness norms for a large set of words; 2) test the extent to which socialness is reliably perceived as a broad construct; 3) explore to what extent socialness captures a distinct aspect of word meaning compared to those measured by other lexical and semantic variables, and 4) assess whether socialness can account for variance in behavioural responses in lexical tasks. To this end, we compiled the largest set of socialness norms available to date by collecting ratings for a set of 8388 English words, including nouns, verbs and adjectives. The socialness ratings show high reliability, and this suggests that the construct is meaningful to participants even at the broad and inclusive level of description provided. Moreover, the validity of the socialness construct was confirmed by a strong correlation with ratings on two other social semantic dimensions (Binder et al., 2016; Troche et al., 2017), despite the distinct definitions employed. However, our socialness measure shared around 58% of its variance with each of these other ratings, possibly reflecting differences in participant characteristics or perhaps methodological choices such as our more inclusive definition which might capture some additional aspects of social experience. Subsequent research will be needed to more thoroughly explore the precise aspects of our interactions with the social environment that are captured by this inclusive socialness measure, such as those measured by more restricted definitions (for examples, see Pexman et al., 2021).

Our preliminary analyses provide some important initial insights into the nature of the socialness dimension. First, while low socialness words tend to be concrete, high socialness words span the entire concreteness continuum, from concrete concepts like *mother*, to more abstract ones like *political*. In line with previous reports of a negative association between a social interaction measure and modality-specific perceptual ratings (Troche et al., 2017), we found that words high in socialness tend to be more abstract and to rely less on sensorimotor information. However, the present findings further suggest that socialness does not relate to concreteness in a simple linear fashion. Although theories of conceptual representation have proposed that social concepts are a sub-type of abstract concepts (Borghi et al., 2019; Kiefer & Harpaintner, 2020), this finding highlights the need to better understand the contribution made by socialness beyond this extreme of the concreteness dimension. Second, we found that words with increased socialness tend to be more valenced and arousing. This is in line with findings that social and affective dimensions reduce to the same latent factor of a multidimensional semantic space (Troche et al., 2014, 2017; Villani et al., 2019). Importantly, while the socialness ratings are significantly correlated with all the lexical and semantic variables explored here, the associated effect sizes are modest and suggest that the socialness

measure captures a distinct aspect of word meaning. This is consistent with fMRI studies which found that the effect of socialness on brain activation during lexical-semantic processing is independent from that of a number of key semantic variables, namely imageability, concreteness, and valence, and suggest that socialness makes a unique contribution to semantic representation (Lin et al., 2018; Wang et al., 2019).

Using regression analyses, we have demonstrated for the first time that socialness of word meaning is related to performance in lexical tasks. This is true even at the broad and inclusive level of description provided. Specifically, we found a facilitatory effect on behavioural performance in lexical decision and word knowledge tasks – increased socialness was associated with faster decision latencies in both tasks and with better accuracy in the LDT. Importantly, this was true after controlling for other semantic variables known to influence lexical-semantic processing, namely concreteness, valence and semantic diversity. Further, this was true even in lexical tasks that involve only shallow semantic processing, where there is a limited pool of variance to be explained by semantic predictors. This unique contribution of the socialness measure suggests that it captures important information about semantic representation and processing and is in line with previous research on semantic richness effects. Semantic richness refers to the phenomenon whereby responses to words that are associated with relatively more semantic information tend to be facilitated in lexical and semantic tasks by virtue of their richer representations that allow faster and more accurate retrieval of meaning (for a review, see Pexman, 2012). As such, increased socialness might enrich a word's conceptual representation and, consequently, facilitate lexical decisions via stronger feedback from semantic to orthographic representations (Hino et al., 2002; Hino & Lupker, 1996). Furthermore, our results suggest that socialness contributes to processing alongside other meaning dimensions derived from multiple experiential channels including linguistic (i.e., semantic diversity), sensorimotor (i.e., concreteness) and affective experience (i.e., valence). This is consistent with theories claiming that conceptual representation is multidimensional in nature and that social experience may be one of the underlying semantic dimensions (e.g., Borghi et al., 2019).

The ability of the semantic dimensions to explain variance in behavioural responses varied depending on the requirements of the task. While socialness and semantic diversity had a facilitatory effect on RTs in both tasks, concreteness and valence contributed to the word knowledge task, but not to the LDT. This is in line with research suggesting that conceptual representations are not stable across time and contexts; instead, the aspects of a word's conceptual representation retrieved at any one point depend on the specific task/context

(Pexman, 2020; Yee & Thompson-Schill, 2016). Our pattern of findings may be explained by the fact that LDT only requires the retrieval of some indication that a word has meaning, such as that indexed by its association with a multiplicity of meanings (i.e., semantic diversity). In comparison, the word recognition task might require access to additional features of a word's meaning, like those that tap into the richness of associated sensorimotor (i.e., concreteness) and emotional experience (i.e., valence extremity). It might also suggest that socialness does not contribute additional semantic features to enrich a word's conceptual representation, but is more indicative of the general relevance or salience of its meaning. This might be consistent with our finding that the socialness of a word does not account for variance in the number of people who know its meaning. Relatedly, it has been observed that social stimuli are preferentially processed during free viewing of complex naturalistic scenes, to the extent that socialness competes with the physical saliency of stimuli (End & Gamer, 2017, 2019). However, future research is needed to better understand the nature of the contribution made by socialness to the semantic richness of concepts (see Muraki et al. (2019) for an example of how to approach examining the factor structure of semantic richness). Moreover, it is important to highlight that, while the words we encounter are typically embedded in rich linguistic contexts (e.g., sentences) that shape our understanding of individual words, the socialness ratings were generated based on words presented in isolation. Future research should address this limitation by moving away from single word processing and considering the lexical-semantic properties of connected text/speech.

5. Conclusions

In the present study, we compiled the largest set of openly-available socialness norms to date. We used an inclusive definition, found that it produced reliable ratings and, thereby, showed that socialness has meaning as a broad construct. An important avenue for future research is identifying the specific aspects of social experience that are most related to conceptual processing to refine our working definition of socialness. Further, our explorations suggest that socialness captures an aspect of word meaning that is distinct to those measured by other key semantic variables and notably, an aspect of meaning that is behaviourally relevant. Our study also provides some initial insights into the information captured by the socialness measure, but subsequent work will be needed on this matter, as well as its role and behavioural consequences across the lifespan, including during acquisition, retrieval and when the semantic system is impaired. Thus, the socialness norms described here will enable future research into the organization and grounding of conceptual knowledge, and can help target testable predictions

about brain and behaviour that can be derived from multiple representation theories (e.g., Borghi et al., 2019) and neurobiological accounts of social semantics (for an extensive discussion, see Pexman et al., 2021; also Binney et al., 2016; Binney & Ramsey, 2020; Diveica et al., 2021).

CHAPTER 4

Establishing a role of the semantic control network in social cognitive processing: a meta-analysis of functional neuroimaging studies

This chapter has been published as:

Authors' contributions and acknowledgements:

Open science practices and data/code availability statement:

The current study was pre-registered via the Open Science Framework; the registration is available at: osf.io/dscwv. The supplementary materials, the raw datasets, as well as the inputs and outputs of all analyses are openly-available at osf.io/fktb8/.

Abstract

The contribution and neural basis of cognitive control is under-specified in many prominent models of socio-cognitive processing. Important outstanding questions include whether there are multiple, distinguishable systems underpinning control and whether control is ubiquitously or selectively engaged across different social behaviours and task demands. Recently, it has been proposed that the regulation of social behaviours could rely on brain regions specialised in the controlled retrieval of semantic information, namely the anterior inferior frontal gyrus (IFG) and posterior middle temporal gyrus. Accordingly, we investigated for the first time whether the neural activation commonly found in social functional neuroimaging studies extends to these 'semantic control' regions. We conducted five coordinate-based meta-analyses to combine results of 499 fMRI/PET experiments and identified the brain regions consistently involved in semantic control, as well as four social abilities: theory of mind, trait inference, empathy and moral reasoning. This allowed an unprecedented parallel review of the neural networks associated with each of these cognitive domains. The results confirmed that the anterior left IFG region involved in semantic control is reliably engaged in all four social domains. This supports the hypothesis that social cognition is partly regulated by the neurocognitive system underpinning semantic control.

CHAPTER 4. SEMANTIC CONTROL IN SOCIAL COGNITION

1. Introduction

The ability to comprehend and respond appropriately to the behaviour of others is essential for humans to survive and thrive. A major challenge for the cognitive sciences, therefore, is to characterise *how* we understand others and coordinate our behaviour to achieve mutually beneficial outcomes, and what can cause this ability to break down (Frith, 2007). There is an indubitable requirement for systems that control, or regulate, the cognitive processes underpinning social interactions. This is because social interactions are intricate and fraught with the potential for misunderstandings and faux pas; first, the everyday social signals to which we are exposed are typically complex, often ambiguous and sometimes conflicting. This is compounded by the fact that the meaning of a given gesture, expression or utterance can vary across contexts (Barrett et al., 2011; Rodd, 2020). Moreover, once we have settled upon an interpretation of these signals, we are then faced with the additional challenge of selecting an appropriate response, and inhibiting others which might, for example, be utilitarian but socially insensitive or even damaging. In order to undergo social interactions that are coherent, effective and context-appropriate, we must carefully regulate both our comprehension of, and response to, the intentions and actions of others (Binney & Ramsey, 2020; Fujita et al., 2014; Gilbert & Burgess, 2008; Ramsey & Ward, 2020).

Despite there being a wealth of literature describing executive functions involved in general cognition (Assem et al., 2020; Diamond, 2013; Duncan, 2010, 2013; Fedorenko et al., 2013; Petersen & Posner, 2012), prominent models of socio-cognitive processing are under-specified in terms of the contribution and neural basis of cognitive control mechanisms (Adolphs, 2009, 2010; Frith & Frith, 2012; Lieberman, 2007). For example, Adolphs (2010) only very briefly refers to cognitive processes involved in 'social regulation' and largely within the limited context of emotional regulation. Likewise, Frith and Frith (2012) refer to a "supervisory system" which has the characteristic features of executive control, but its functional and anatomical descriptions lack detail important for generating testable hypotheses. However, research into specific social phenomena, such as prejudice (Amodio, 2014; Amodio & Cikara, 2021) and automatic imitation (Cross et al., 2013; Darda & Ramsey, 2019) has recently begun to give the matter of cognitive control greater attention. Of particular interest has been the contribution of the domain-general multiple-demand network (MDN), a set of brain areas engaged by cognitively-challenging tasks irrespective of the cognitive domain (Assem et al., 2020; Duncan, 2010; Fedorenko et al., 2013; Hugdahl et al., 2015). MDN activity increases with many kinds of general task demand, including working memory load and task

switching, and it has been suggested that this reflects the implementation of top-down attentional control and the optimal allocation of cognitive resources to meet immediate goals (Duncan, 2010, 2013). The MDN is comprised of parts of the precentral gyrus, the middle frontal gyrus (MFG), the intraparietal sulcus (IPS), insular cortex, the pre-supplementary motor area (pre-SMA) and the adjacent cingulate cortex (Assem et al., 2020; Fedorenko et al., 2013), some of which have been implicated in controlled social processing such as, for example, working memory for social content (Meyer et al., 2012), social conflict resolution (Zaki et al., 2010), inhibition of automatic imitation (Darda & Ramsey, 2019) and mental state inference or theory of mind (ToM) (Rothmayr et al., 2011; Samson et al., 2005; Van der Meer et al., 2011). However, there are at least three key unresolved questions regarding the role of cognitive control in social cognition. First, it remains to be seen whether there could be multiple, distinguishable types of, and neural systems for, control. Second, it is unclear whether distinguishable control systems are necessary for all or only certain social abilities and, third, whether this engagement depends on specific task demands. Shedding light on these issues has the potential to generate important new hypotheses regarding social behaviour both in the context of health and injury/disease.

It has recently been proposed that a relatively specialised form of cognitive control, termed *semantic control*, could be particularly important for social cognitive processing (Binney & Ramsey, 2020). This proposal argued that a semantic control system is required during social cognitive tasks to modulate the retrieval and selection of conceptual-level knowledge so that it is relevant to the situational context or the task at hand (Chiou et al., 2018; Jefferies, 2013; Lambon Ralph et al., 2017). The reasons why semantic control should be critical for social cognition are uncomplicated; we retain a vast amount of socially-relevant knowledge including knowledge about familiar people (Greven et al., 2016; Hassabis et al., 2014), about the structure of and relationship between social categories and their associated stereotypes (Freeman & Johnson, 2016; Quinn & Rosenthal, 2012), and an understanding of abstract social concepts, norms and scripts (Frith & Frith, 2003; Van Overwalle, 2009). But only a limited portion of this information is relevant in a given social instance and it would be computationally inefficient to automatically retrieve it all. For example, there is no need to retrieve information about someone's personality traits, or personal interests and hobbies, if the only task is to pick them out from within a crowd. Moreover, the types and the scope of information we need to retrieve to understand and respond appropriately to certain social signals change according to the context, and irrelevant information could potentially interfere. Therefore, semantic control should be essential for limiting potential social errors.

There is a growing body of convergent patient, neuroimaging and neuromodulation evidence that semantic memory retrieval engages the semantic control network (SCN) which comprises the anterior IFG and the posterior middle temporal gyrus (pMTG) (Badre et al., 2005; Jackson, 2021; Noonan et al., 2010; Whitney et al., 2012). While the domain-general MDN is also engaged by semantic tasks, and particularly those with high control demands (Jackson, 2021; Thompson et al., 2018), there is evidence to suggest that both the anatomy of the SCN and MDN and their functional contributions to controlled semantic processing are distinct (Gao et al., 2021; Humphreys et al., 2017; Wang et al., 2018; Whitney et al., 2012). In particular, fMRI studies revealed that the mid- to posterior IFG (pars triangularis and pars opercularis), nodes of the MDN, increase activity in response to increased 'semantic selection' demands, a process that is engaged when automatic retrieval of semantic knowledge results in competition between multiple representations which must be resolved (for example, hearing the word *bank* might elicit retrieval of the concept of a riverside and a financial institution) (Badre et al., 2005; Nagel et al., 2008; Thompson-Schill et al., 1997). However, this mid- to posterior IFG region is also engaged by other non-semantic forms of response competition (Badre & Wagner, 2007; Dobbins & Wagner, 2005) and tests of inhibitory function such as the Stroop task (Huang et al., 2020; January et al., 2009; Nee et al., 2007). In contrast, activation of the anterior IFG (pars orbitalis) appears to be more selective to semantic control demands and driven specifically by an increased need for 'controlled semantic retrieval', a mechanism that is engaged when automatic semantic retrieval fails to activate semantic information necessary for the task at hand, and a further goal-directed semantic search needs to be initiated (Gold et al., 2006; Krieger-Redwood et al., 2015; Wagner et al., 2001).

To date, there have been but a few neuroimaging investigations that have directly questioned the involvement of the SCN in social cognitive processing. Two recent fMRI studies compared activation during semantic judgements made on social and non-social stimuli and found that the IFG and pMTG were engaged by both stimulus types (Binney, Hoffman, et al., 2016; Rice et al., 2018). Further, Satpute et al. (2014) found that controlled retrieval, but not selection of social conceptual information engages the anterior IFG. However, we are not aware of any prior studies that attempt to examine the engagement of the SCN specifically during tasks that are commonly viewed as social in nature (e.g., ToM tasks). As a starting point, rather than conducting a novel individual experiment, the present study adopted a meta-analytic approach to extract reliable trends from large numbers of studies. Meta-analyses of functional neuroimaging data overcome the limitations of individual studies (Cumming, 2014; Eickhoff et al., 2012), which are frequently statistically underpowered (Button et al., 2013) and

vulnerable to effects of idiosyncratic design and analytic choices (Botvinik-Nezer et al., 2020; Carp, 2012) so that it becomes difficult to distinguish between replicable and spurious findings and to generalize the results. Our principal aim was to determine whether the distributed neural activation commonly associated with functional neuroimaging studies of social cognition extends to the neural networks underpinning semantic control (i.e., SCN and MDN). In order to localise the brain network sensitive to semantic control demands (i.e., semantic retrieval and/or selection), and then compare and contrast it to networks implicated in social cognition, we performed an update of Noonan et al.'s (2013) meta-analysis of semantic control (also see Jackson, 2021).

We took the approach of investigating multiple sub-domains of social cognition in parallel because this should allow an assessment of the extent to which inferences are generalisable, rather than specific to certain types of social tasks and/or abilities. We chose to focus on four particular areas of research that target abilities frequently identified as key facets of the human social repertoire - ToM, empathy, trait inference, and moral reasoning (Lieberman, 2007; Van Overwalle, 2009) – and, for each, we conducted separate meta-analyses of the available functional imaging data to determine the brain regions consistently implicated. In the case of trait inference, this was the first neuroimaging meta-analysis to include studies that used stimuli other than faces (see Section 2, and also Bzdok et al., 2011, and Mende-Siedecki et al., 2013, for contrasting approaches). In the other three cases, we performed updates of prior meta-analyses (Eres et al., 2018; Molenberghs et al., 2016; Timmers et al., 2018).

Further, we conducted an exploratory conjunction analysis aimed at identifying brain areas reliably implicated in all four social sub-domains and, thus, a core network for social cognitive processing (Bzdok et al., 2012; Schurz et al., 2020; Van Overwalle, 2009). We hypothesised that this core network would include parts of the MDN and the SCN. It is of note that, across all four social sub-domains, we took a different approach to study inclusion and exclusion criteria than that taken by some prior meta-analyses of general social cognition (Van Overwalle, 2009). In particular, we excluded studies investigating processes associated primarily with the self because social cognition is, although perhaps only in the strictest sense, about understanding other people. We also excluded studies in which tasks could be completed based on relatively simple perceptual processing and without a need for deeper cognitive and inferential processes (e.g., emotion discrimination tasks, automatic imitation). This was done in an attempt to constrain our inferences to be about the neurobiology underpinning cognitive

rather than primarily perceptual social processes (for further detail on this distinction, see Adolphs, 2010; and Binney & Ramsey, 2020).

Finally, as a secondary aim, the present study used the meta-analytic approach to assess whether there are differences in the neural networks engaged by implicit and explicit social processing (also see Dricu & Frühholz, 2016; Eres et al., 2018; Fan et al., 2011; Molenberghs et al., 2016; Timmers et al., 2018; Van Overwalle & Vandekerckhove, 2013). This was aimed at addressing a pervasive distinction in the social neuroscientific literature between automatic and controlled processes (Adolphs, 2010; Happé et al., 2017; Lieberman, 2007), and followed an assumption that implicit paradigms engage only automatic processes, whereas controlled processes are recruited during explicit paradigms (Sherman et al., 2014). Automatic processes are described as unintentional, effortless, and fast, whereas controlled processes are deliberate, effortful, and thus slower (Lieberman, 2007; Shiffrin & Schneider, 1977). Some authors have argued that automatic and controlled social processes are mutually exclusive of one another and draw upon distinct cortical networks, with the former engaging lateral temporal cortex, the amygdala, ventromedial frontal cortex and the anterior cingulate, and the latter engaging lateral and medial prefrontal and parietal cortex (Forbes & Grafman, 2013; Lieberman, 2007). However, these dual-process models have been criticised for over-simplifying both the distinction and the relationship between automatic and controlled processes (Amodio, 2019; Cunningham & Zelazo, 2007; Ferguson et al., 2014; Fidler & Hütter, 2014; Fujita et al., 2014; Melnikoff & Bargh, 2018). An alternative proposal, that we describe above, makes a different distinction - one between representation and control. This neurocognitive model proposes that social processing relies on a single-route architecture wherein the degree to which cognitive processing has certain attributes (e.g., speed or effort) does not reflect one system versus another. Instead, it is proposed that it reflects the degree to which the control system needs to exert influence, upon otherwise automatic activation within the representational system, in order to meet the demands of a task in an appropriate and efficient manner (Binney & Ramsey, 2020; Jefferies, 2013). If the dual route model is correct, explicit but not implicit social paradigms should differentially engage brain regions associated with cognitive control demands, including the SCN and MDN. If the single-route model is correct, then there should be no qualitative difference in terms of the network of regions activated by implicit paradigms (ergo automatic processing) and explicit paradigms (ergo controlled processing), although there may be differences in the magnitude of regional activation.

To summarise, the aims of the present study were as follows: 1) explore the involvement of domain-general control systems in social cognition; more specifically,

determine whether social cognitive processing reliably engages brain areas implicated in the controlled retrieval and selection of conceptual knowledge; and 2) examine the evidence for dual-route and single-route models of controlled social cognition.

2. Methods

Following open science initiatives (Munafò et al., 2017), the current study was pre-registered via the Open Science Framework (OSF). We adhered to our pre-registered protocols (available at: osf.io/dscwv) with a few minor exceptions (see Section S1 of Supplementary Information (SI) 1 for details). All the raw datasets are openly-available on the OSF project page (available at: osf.io/fktb8/) and are accompanied by a range of study characteristics including details that are not the focus of the present study but may be of interest in future research (please see Section S1 of SI1 for a detailed description). Moreover, the input data and output files of all analyses can be accessed via the OSF page.

In accordance with our pre-registered aims, we performed a comprehensive review of published functional neuroimaging studies investigating four social abilities – Theory of mind (ToM), trait inference, empathy and moral reasoning - and independent coordinate-based meta-analyses aimed at characterising the brain-wide neural networks underpinning each. In the case of three of these domains (ToM, empathy and moral reasoning), we updated earlier meta-analyses (Eres et al., 2018; Molenberghs et al., 2016; Timmers et al., 2018), capitalizing on additional data, and also implementing recommendations for best practice that became available in a year subsequent to these prior studies (Müller et al., 2018). In the case of trait inference, as far as we are aware, this was the first neuroimaging meta-analysis to include studies that explored potential sources of information beyond face stimuli (for contrasting approaches see Bzdok et al., 2011; Mende-siedlecki et al., 2013). To localise the brain areas underpinning semantic retrieval and selection, we also updated a meta-analysis of functional imaging studies of semantic control by Noonan et al. (2013). This involved the inclusion of additional data, and improvements in meta-analytic tools which corrected previous implementation errors that led to the use of liberal statistical thresholds (Eickhoff et al., 2017).

To directly address our first aim, we assessed the degree of overlap between the neural networks supporting semantic control and those involved in social information processing via a set of formal conjunctions and contrasts analyses. To address our second aim, where possible, we contrasted brain-wide activation associated with explicit versus implicit social cognitive paradigms. Tasks that drew the participant's attention to the behaviour/cognitive process of interest were categorised as explicit, while tasks that used non-specific instructions (e.g., they

involved passive observation of stimuli) or employed orthogonal tasks (e.g., age judgement) were categorised as implicit. Finally, where sufficient relevant information was available, we explored the influence of task difficulty on patterns of brain activation.

All of the meta-analyses reported below were conducted following best-practice guidelines recommended by Müller et al. (2018). This, as well as several refinements to inclusion/exclusion criteria, contributed to methodological differences between the present meta-analyses and those prior meta-analyses upon which the 'updates' were based. A summary of similarities and differences is provided in Table S1 (SI1) and further details are given in the sections below.

2.1. Literature selection and inclusion criteria

2.1.1. General approach and criteria

Where possible, relevant functional neuroimaging studies were initially identified based on their inclusion in a recent prior neuroimaging meta-analysis. These lists were supplemented via a search on the Web of Science (WoS) online database (www.webofknowledge.com) for original reports published in the years subsequent, and by searching through reference lists of said articles. Each WoS search used the terms ['fMRI' or 'PET'], as well as terms uniquely chosen for a given cognitive domain (see Table 1).

Table 1. Terms used to search the Web of Science database for relevant articles.

Cognitive domain	Search terms
Semantic control	'semantic', 'comprehension', 'conceptual knowledge', 'selection', 'retrieval', 'inhibition', 'control', 'elaboration', 'fluency', 'ambiguity', 'metaphor', 'idiom'
Theory of Mind	'theory of mind', 'ToM', 'mentalising', 'mentalizing'
Trait inference	'social judgement', 'social evaluation', 'social attribution', 'trait inference', 'impression formation'
Empathy	'empathy', plus 'empath*' - corresponding variations (e.g. 'empathic')
Moral cognition	'morality', 'moral', 'moral decision making', 'moral emotion', 'harm', 'guilt'

N.b., For all five cognitive domains, the search followed the following format: [fMRI OR PET] AND [term1 OR term2 OR … OR termX].

A general set of inclusion criteria applied to all our analyses were as follows:

1) Only studies that employed task-based fMRI or PET to obtain original data were included. Studies employing other techniques (e.g., EEG/MEG), meta-analyses and review articles were excluded.

2) Studies were only included if they reported whole-brain activation coordinates that were localised in one of two standardised spaces – Talairach (TAL) or Montreal Neurological Institute (MNI) – or these coordinates were made available on request (see Section 1 of SI1). Coordinates reported in TAL space were converted into MNI space using the Lancaster transform (tal2icbm transform (Lancaster et al., 2007) embedded within the GingerALE software (version 3.0.2; http://brainmap.org/ale). Studies exclusively reporting results from region-of-interest or small volume correction analyses were excluded because these types of analysis violate a key assumption of coordinate-based meta-analyses (Eickhoff et al., 2012; Müller et al., 2018).

3) Studies were only included if they reported activation coordinates that resulted from univariate contrasts clearly aimed at identifying the process of interest (e.g., ToM). We included contrasts between an experimental task and either a comparable active control task or a low-level baseline such as rest or passive fixation. Contrasts against low-level baselines were included in the primary analyses because they can reveal activity associated with domain-general cognitive processes that is subtracted out by contrasts between active conditions. This could include semantic processes that are common to both social and non-social tasks. However, contrasts against low-level baselines also yield activity associated with differences in perceptual stimulation and attentional demand. To address this caveat, we repeated the analyses whilst excluding this subset of contrasts. The results can be found on the project's OSF page (available at: osf.io/fktb8/). We excluded contrasts that make comparisons between components of the process of interest (e.g., affective vs. cognitive ToM; utilitarian vs. deontological moral judgements) because we were interested in the common, core processes that would be subtracted out by these contrasts (but see the following paragraph).

4) Multiple contrasts from a single group of participants (e.g., separate contrasts against one of two different baseline conditions) were included in a single meta-analysis as long as they independently met all other inclusion criteria for the primary analyses. This allowed maximum use of all available data and enabled us to evaluate the effect of using different types of baseline, for example (see above). However, it is important to

adjust for this (Müller et al., 2018), and accordingly, we adopted an approach to controlling for within-group effects (Turkeltaub et al., 2012); specifically, sets of activation coordinates from different contrasts, but the same participant group, were pooled. This means that when we refer to the numbers of experiments, we have counted multiple contrasts from a single participant sample as one single experiment. In cases where two or more published articles contained data from the same participant sample, we pooled distinct contrasts as above, and excluded duplicates. This partially explains why the number of experiments in our analyses is lower than in those of some prior meta-analyses. However, in formal contrast analyses that compare different conditions (e.g., instructional cue, task difficulty), contrasts like these would be separated, and care was also taken to minimize the difference in the number of experiments on either side of the contrast. For example, if a study reported two contrasts – one implicit and one explicit - based on the same participant group, only the peaks from the implicit task would be included in the contrast/conjunction analyses if there were a greater number of explicit than implicit tasks overall (see Figure S9 of SI1).

5) Only studies that tested healthy participants were included. Contrasts including clinical populations or pharmacological interventions were excluded.

6) Only research articles published in English were included.

2.1.2. Theory of mind

This meta-analysis was built upon that of Molenberghs et al. (2016) and only included studies that were specifically designed to identify the neural network underpinning ToM processes (i.e., they employed tasks involving inferences about the mental states of others, including their beliefs, intentions, and desires). Therefore, studies that looked at passive observation of actions, social understanding, mimicry or imitation were not included, unless tasks included a ToM component. Unlike Molenberghs et al. (2016), we excluded studies investigating irony comprehension (e.g., Wang et al., 2006) because ToM might not always be necessary to detect non-literal meaning in language (Ackerman, 1983; Bosco et al., 2018; Pexman, 2008) and studies that employed interactive games (e.g., Rilling et al., 2008). These latter studies are commonly designed to investigate the degree to which ToM is engaged under different task conditions rather than distinguish activation associated with ToM from that related to other processes. Moreover, unlike Molenberghs et al. (2016), we excluded studies that employed trait inference tasks as these were considered separately (see Section 2.1.3).

The Molenberghs et al. (2016) search was inclusive of fMRI studies published prior to July 2014 and yielded 144 independent experiments (1789 peaks) contributing to their analysis. We performed a WoS search for further original fMRI and PET studies conducted between August 2014 and March 2020, and a search of PET studies published prior to July 2014. We then applied our inclusion criteria to both newly identified studies and those analysed by Molenberghs and colleagues (see Table S1 of SI1 for further differences in criteria). In the end, we found 136 experiments with a total number of 2158 peaks and 3452 participants that met our criteria for inclusion (see Figure S1of SI1 for more details regarding the literature selection process; and Table S1 of SI2 for a full list of the included experiments).

2.1.3. Trait inference

Studies were included in the meta-analysis if they employed tasks that required the participants to infer the personality traits of others based on prior person knowledge or another's appearance and/or behaviour. Whereas the types of mental states typically inferred in ToM tasks are transitory in nature (e.g., relating to immediate goals or the intentions behind a specific instance of behaviour), traits are coherent and enduring dispositional characteristics of others (i.e., personality traits; Van Overwalle, 2009). Previous meta-analyses (Molenberghs et al., 2016; Schurz et al., 2014) of ToM have included tasks requiring trait inferences. However, it has been suggested that personality trait inferences differ from mental state inferences in terms of likelihood and speed of processing, and hold a higher position in the hierarchical organisation of social inferential processes (Korman & Malle, 2016; Malle & Holbrook, 2012). In line with this proposal, we maintained a distinction and performed separate analyses. Moreover, previous imaging meta-analyses of trait inference were limited to studies that used face stimuli (Bzdok et al., 2011; Mende-siedlecki et al., 2013). However, trait inferences can be made on the basis of many different sources of information, including physical appearance, behaviour and prior knowledge about others (Uleman et al., 2007). To our knowledge, the present attempt is the first to include studies that required participants to make trait inferences based on facial photographs, behavioural descriptions *or* prior person knowledge. We excluded any studies that asked participants to make inferences about transitory mental states, including basic emotions. We also excluded studies that did not use a subtraction approach, but rather investigated brain activity that varied parametrically with the levels of a pre-defined trait dimension (e.g. Engell et al., 2007). Finally, we excluded studies that included emotional face stimuli to avoid conflating brain activity related to trait inference with that associated with emotion recognition and processing.

We performed a WoS search of studies published before March 2020 and reference-tracing to identify relevant studies for inclusion in the meta-analysis. A total of 40 experiments with 523 peaks and 732 participants were found to meet the criteria for inclusion (Figure S2 of SI1; Table S2 of SI2).

2.1.4. Empathy

This meta-analysis was built upon that of Timmers et al., (2018) and only included studies that were specifically designed to identify the neural network underpinning empathy by employing tasks asking participants to observe, imagine, share and/or evaluate the emotional or sensory state of others. The task definition was kept identical to previous meta-analyses on empathy (Fan et al., 2011; Timmers et al., 2018). We also made a distinction between tasks eliciting empathic responses to other people's pain and those investigating empathic responses to others' affective states.

Timmers et al., (2018) included studies published before December 2017, totalling 128 studies with 179 contrasts (1963 peaks). We identified additional original studies conducted between January 2018 and March 2020 via a WoS search and subsequently applied our inclusion criteria to all, including those analysed by Timmers et al. (2018) (see Table S1 of SI1 for further differences in criteria). This resulted in a yield of 163 experiments with a total number of 2691 peaks and 4406 participants (Figure S3 of SI1; Table S3 of SI2). Empathy for pain was independently investigated in 93 of these experiments, empathy for affective states was independently explored in 69 experiments, and 9 experiments concurrently explored both empathy for pain and emotions in the same contrasts.

2.1.5. Moral reasoning

This analysis updated a previous meta-analysis conducted by Eres et al. (2018) and included studies that employed tasks designed to investigate judgements and decision-making based on moral values. In line with Eres et al. (2018), studies that did not specifically have a morality component were not included. For example, studies investigating judgements regarding adherence to social expectations but not moral values (e.g., Bas-Hoogendam et al., 2017) were excluded.

Eres et al. (2018)'s search was restricted to fMRI studies and covered the period before February 2016 yielding 123 contrasts (989 peaks). We expanded this list via a WoS search for original fMRI and PET studies published between March 2016 and March 2020, and a search for PET studies published before March 2016, and then applied our inclusion criteria (see Table

S1 of SI1 for differences in criteria). This resulted in a yield of 68 experiments with a total number of 884 foci and 1587 participants (Figure S4 of SI1; Table S4 of SI2).

2.1.6. Semantic control

In this meta-analysis, we sought to extend an earlier meta-analysis conducted by Noonan et al. (2013; also see Jackson, 2021). In line with theirs, this analysis only included studies that were specifically investigating semantic processing, and that reported contrasts that reflected high > low semantic control within a semantic task, or comparisons between a task requiring semantic control and an equally demanding executive decision in a non-semantic domain. We excluded studies with a focus upon priming without an explicit semantic judgement (e.g., primed lexical decision), bilingualism, episodic memory, or sleep consolidation.

Noonan et al. (2013)'s search covered the period between January 1994 and August 2009 and yielded 53 studies (395 peaks) that met their criteria for inclusion in their analysis. We performed a WoS search for original studies published between September 2009 and March 2020, and reference-tracing, and then applied our inclusion criteria to both newly identified studies and those analysed by Noonan et al. (2013). This produced a yield of 92 experiments with a total number of 971 peaks and 1966 participants that met the criteria for inclusion in our analysis (Figure S5 of SI1; Table S5 of SI2).

2.2.   Data analysis

We performed coordinate-based meta-analyses using the revised activation likelihood estimation (ALE) algorithm (Eickhoff et al., 2009, 2012; Turkeltaub et al., 2012) implemented in the GingerALE 3.0.2 software (http://brainmap.org/ale). We used the GingerALE software to conduct two types of analysis. The first were independent dataset analyses, which were used to identify areas of consistent activation across particular sets of experiments. These analyses were performed only on the experiment samples with a recommended minimum of 17 experiments in order to have sufficient power to detect consistent effects and circumvent the possibility of results being driven by single experiments (Eickhoff et al., 2016). The ALE meta-analytic method treats reported activation coordinates as the centre points of three-dimensional Gaussian probability distributions which take into account the sample size (Eickhoff et al., 2009). First, the spatial probability distributions of all coordinates reported were aggregated, creating a voxel-wise modelled activation (MA) map for each experiment. Then, the voxel-wise union across the MA maps of all included experiments was computed, resulting in an ALE map that quantifies the convergence of results across experiments (Turkeltaub et al.,

2012).The version of GingerALE used in the present study tests for above-chance convergence between experiments (Eickhoff et al., 2012) thus permitting random-effects inferences.

Following the recommendations of Müller et al. (2018), for the main statistical inferences, the individual ALE maps were thresholded using cluster-level family-wise error (FWE) correction of $p < .05$ with a prior cluster-forming threshold of $p < .001$ (uncorrected). Cluster-level FWE correction has been shown to offer the best compromise between sensitivity to detect true convergence and spatial specificity (Eickhoff et al., 2016). However, we subsequently applied an additional and more conservative threshold at the voxel level (FWE corrected at $p < .05$). This level of thresholding suffers from decreased sensitivity to true effects, but has the advantage of allowing an attribution of significance to each voxel and thereby increases the spatial specificity of inferences (Eickhoff et al., 2016). The FWE-corrected cluster-level and voxel-height thresholds were estimated using a permutation approach with 5000 repetitions (Eickhoff et al., 2012). None of the meta-analyses that we updated had used the recommended cluster-level FWE or the FWE height-based correction methods.

The second set of analyses, conjunction and contrast analyses, were also performed in GingerALE and were aimed at identifying similarities and differences in neural activation between the different sets of studies. The conjunction images were generated using the voxel-wise minimum value (Nichols et al., 2005) of the included ALE maps to highlight shared activation. Contrast images were created by directly subtracting one ALE map from the other to highlight unique neural activation associated with each dataset (Eickhoff et al., 2011). Then, the differences in ALE scores were compared to a null-distribution estimated via a permutation approach with 5000 repetitions. The contrast maps were thresholded using an uncorrected cluster-forming threshold of $p < 0.001$ and a minimum cluster size of 200 mm^3.

In addition, we performed post-hoc analyses to investigate if the clusters of convergence revealed by the ALE analyses were driven by experiments featuring specific characteristics of interest (i.e., type of instructional cue, task difficulty). To this end, we examined the list of experiments that contributed at least one peak to each ALE cluster and compared the number of contributing experiments featuring the characteristic of interest (e.g., explicit vs implicit processing) by conducting Fisher's exact tests of independence and post-hoc pairwise comparisons (using False Discovery Rate correction for multiple comparisons) in RStudio Version 1.2.5001 (RStudio Team, 2020).

A full list of the confirmatory and exploratory analyses we conducted can be found in Section S3 of SI1.

3. Results

3.1. The 'social brain'

3.1.1. Theory of mind

Convergent activation across all 136 ToM experiments was found in 13 clusters (see Figure 1a and Table S1.1.1 of SI3) located within the bilateral middle temporal gyrus (MTG) (extending anteriorly towards the temporal poles and also in a posterior and superior direction towards the superior temporal gyrus (STG) and angular gyrus (AG) in both hemispheres), bilateral IFG, bilateral dorsal precentral gyrus, ventromedial prefrontal cortex (vmPFC), dorsomedial prefrontal cortex (dmPFC), pre-SMA, precuneus, left fusiform gyrus and left and right cerebellum. All these clusters survived both the height-based and extent-based thresholding. A cluster in the posterior cingulate cortex (PCC) survived height-based thresholding but did not survive extent-based thresholding. These results are largely consistent with those of Molenberghs et al. (2016), with the difference being that they did not find activation in SMA, left fusiform gyrus or cerebellum. In order to address concerns regarding the validity of some other popular ToM tasks (Heyes, 2014; Quesque & Rossetti, 2020), we conducted a separate supplementary meta-analysis that was limited to the subset of ToM experiments that employed false belief tasks (see Section S3.1 of SI1; Table S1.1.2). This analysis revealed convergent activation in similar temporo-parietal and medial frontal regions to the inclusive ToM analysis but did not implicate the lateral frontal cortex.

3.1.2. Trait inference

The ALE meta-analysis revealed convergent activation across 40 experiments in 8 clusters (Figure 1b, Table S1.2) implicating the bilateral IFG, dmPFC, vmPFC, PCC, right pMTG (extending to AG), left AG and left anterior MTG. Voxels from all clusters, except for those in the right pMTG and vmPFC, survived the more conservative height-based thresholding.

3.1.3. Empathy

The ALE meta-analysis of all 163 empathy experiments revealed 16 clusters of convergent activation (Figure S7a; Table S1.3.1), including in the bilateral IFG (extending towards the insula), SMA, dmPFC, bilateral posterior inferior temporal gyrus (ITG), right pMTG, bilateral supramarginal gyrus (SMG), left inferior parietal lobule (IPL), bilateral occipital cortex, bilateral amygdala, left thalamus, left caudate and brainstem. These clusters survived both the height-based and extent-based thresholding, except for the anterior dmPFC and right pMTG clusters which survived extent-based thresholding only. One cluster in the right cerebellum

survived height-based thresholding but did not survive cluster extent-based thresholding. These areas were also implicated by Timmers et al. (2018). In contrast, however, we did not find convergent activation in the left posterior fusiform gyrus, left SMG (although we found a cluster slightly more posterior and inferior), left anterior ITG, right TP, precuneus, middle cingulate gyrus, and right superior parietal lobule.

The separate ALE maps for empathy for pain and empathy for affective states are displayed in Figure 1c and d. A conjunction analysis found activation common to empathy for pain (Table S1.3.2) and empathy for affective states (Table S1.3.3) in the bilateral insula (extending to the IFG), SMA, right precentral gyrus, bilateral ITG, left occipital cortex and the brainstem (Figure S7b; Table S1.3.4). Formal contrasts revealed that empathy for pain and empathy for emotions also engage highly distinct brain areas (Figure S7b; Table S1.3.4). Clusters with increased convergence for empathy for pain were found in left IFG (pars triangularis), left precentral gyrus, bilateral insula, middle cingulate gyrus, bilateral SMG, right IPL and bilateral pITG. In contrast, increased convergence in empathy for affective states was revealed in PCC and right temporal pole. Given these significant differences in their underlying neural networks, empathy for pain and empathy for emotions were considered separately for all subsequent analyses.

3.1.4. Moral reasoning

Convergent activation across all 68 experiments studying moral reasoning was found in 12 clusters (Figure 1e, Table S1.4) located in the left IFG, left insula (extending towards the superior temporal pole), right superior temporal pole (extending towards pars orbitalis of the IFG and insula), mPFC, medial orbitofrontal cortex (OFC), precuneus, bilateral pMTG, and the bilateral anterior MTG. Only four clusters - left insula, mPFC, precuneus and left pMTG - survived height-based thresholding. These results are mostly consistent with those obtained by Eres et al. (2018), with the difference that we did not find convergent activation in the left amygdala and right AG, and found additional clusters of convergent activation in left MFG, bilateral anterior MTG, and right pMTG.
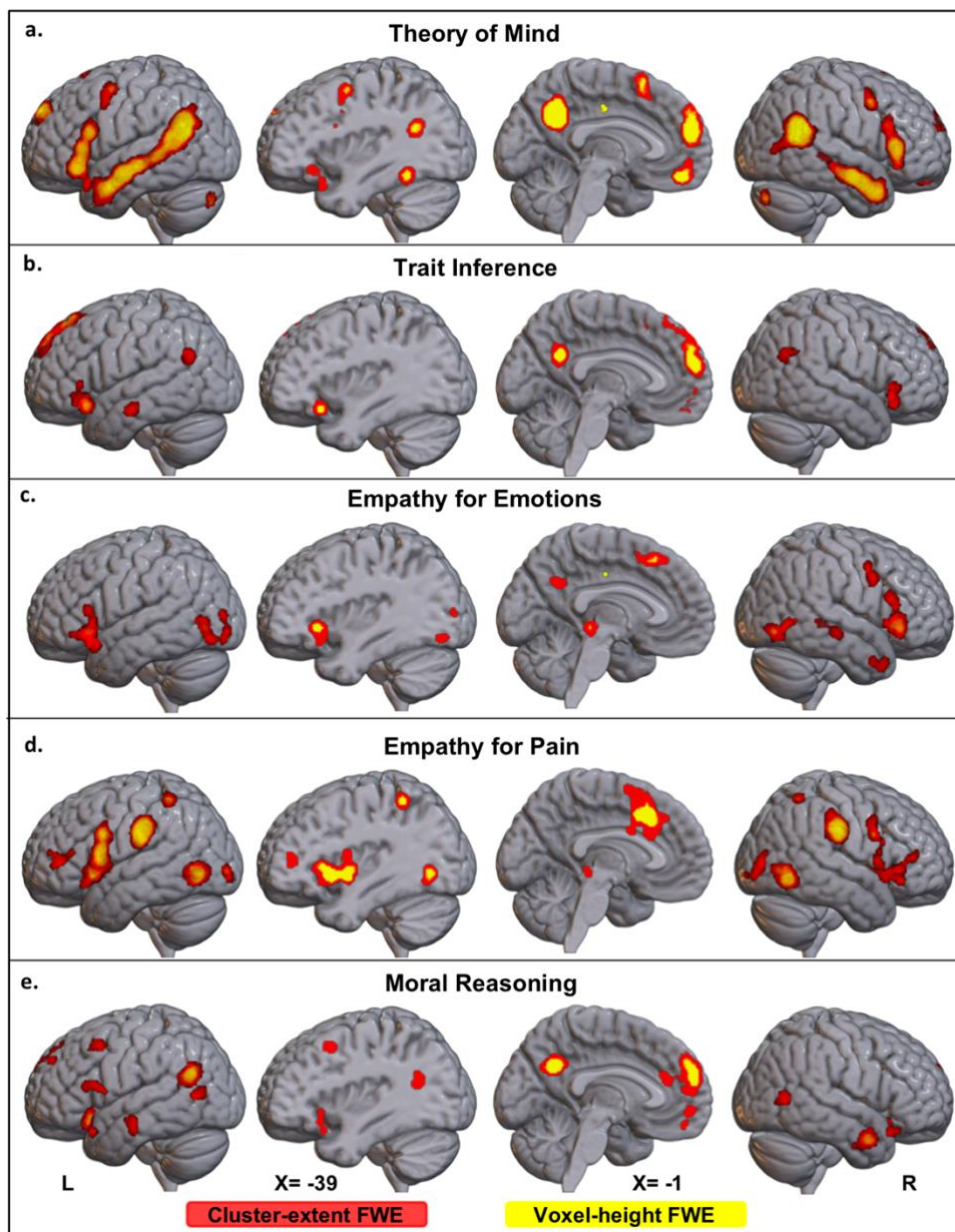
Figure 1. Binary whole-brain ALE maps showing statistically significant convergent activation resulting from independent meta-analyses of ToM studies (N=136), trait inference (N= 40), empathy for pain (N=85) and emotions (N=69) and moral reasoning (N=68). The ALE maps were thresholded using an FWE corrected cluster-extent at $p < .05$ with a cluster-forming threshold of $p < .001$ (red) and an FWE corrected voxel-height threshold of $p < .05$ (yellow). The lateral views, which show projections on the cortical surface, are accompanied by brain slices at the sagittal midline and also coplanar with the peak of the left IFG cluster observed across all social domains (X = -39; Table S1.5).

3.1.5. A common network for multiple sub-domains of social cognition

To identify brain areas consistently activated across multiple sub-domains of social cognition, we performed an overlay conjunction analysis of the cluster-extent FWE-corrected ALE maps associated with ToM, trait inference, empathy (for pain and/or emotions) and moral reasoning (see Figure 2a, Table S1.5). Convergent activation across all four socio-cognitive sub-domains was found in the left IFG (pars orbitalis) and precuneus. Overlapping areas of activation across three of four social sub-domains included right IFG, left IFG (pars triangularis and pars opercularis), SMA, mPFC, medial OFC, left MTG, left pMTG/AG, right anterior MTG and right pMTG/ITG. Overlap between two of four maps was found in bilateral precentral gyrus, right AG, right pMTG and left pMTG/ITG. Because the conservative thresholding used in this analysis could have excluded smaller clusters that nonetheless overlap across the sub-domains, we repeated the conjunction using ALE maps treated with a more liberal statistical threshold of $p<.001$ uncorrected. This revealed additional overlapping activation for all four social domains in the right IFG (pars orbitalis), mPFC, left pSTG/AG and bilateral ATL (Figure S8). These brain areas have been implicated in a variety of social-cognitive abilities by multiple previous meta-analyses (Alcalá-López et al., 2018).

The extent to which brain regions engaged in social cognition overlap with those engaged in general semantic cognition (including both representation and control processes) is illustrated in Figure 2b. Figure 2c shows that the brain regions engaged in social cognition are largely non-overlapping with the network engaged by domain-general executive processes (i.e., the MDN).
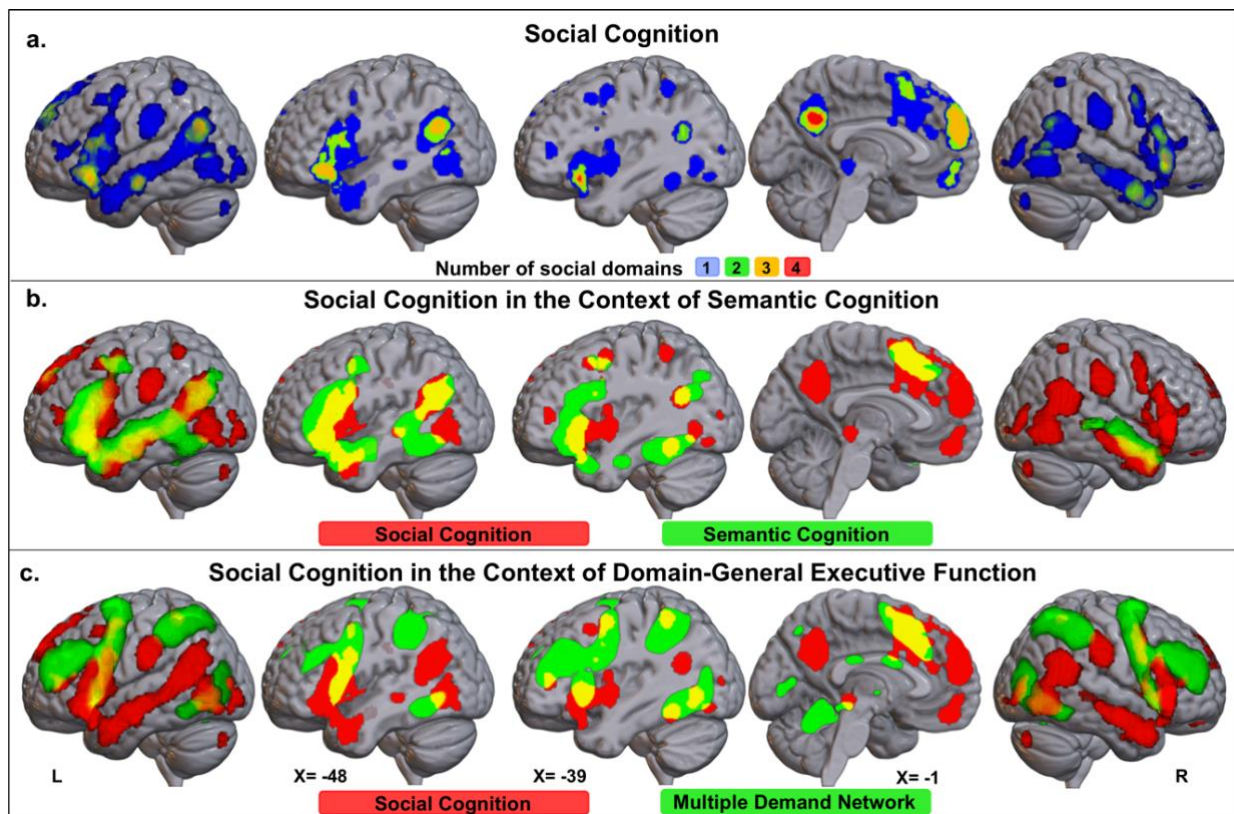
Figure 2. The neural network engaged in social cognitive processing: (a) An overlay conjunction of the ALE maps resulting from independent meta-analyses of ToM studies, trait inference, empathy for pain/emotions, and moral reasoning. The map displays the number of social domains showing convergent activation in each voxel. The ALE maps were thresholded using an FWE corrected cluster-extent threshold at *p* < .05 with a cluster-forming threshold of *p* < .001. (b) The binarized social cognition map (red) generated by the overlay conjunction is displayed overlaid with a binarized ALE map of convergent activation across N = 415 semantic > non-semantic contrasts generated in Jackson, 2021 (green); overlap is shown in yellow. (c) The binarized social cognition map (red) generated by the overlay conjunction is displayed overlaid with a mask of the multiple-demand network (MDN) generated in Federenko et al., 2013 (green) by contrasting hard > easy versions of seven diverse cognitive tasks; overlap is shown in yellow. The lateral views, which show projections on the cortical surface, are accompanied by brain slices at the sagittal midline and also coplanar with the peak of the left STG (X = -48) and left IFG (X = -39) clusters that overlapped across all four social domains (Table S1.5).

3.2. The semantic control network

The ALE meta-analysis of all 92 semantic control experiments revealed convergent activation in a distributed network consisting of frontal, temporal and parietal areas (Figure 3a, Table S2). The largest cluster was located in the left frontal lobe and extended from the IFG (pars orbitalis) to MFG. In the right frontal lobe, convergent activation was limited to two clusters with peaks in pars orbitalis and pars triangularis of the IFG. Consistent activation was also found in the medial frontal cortex with the peak in SMA. The left temporal cluster extended from the posterior portion of the MTG, which showed the highest level of convergence, to the fusiform gyrus. All these clusters survived both the height-based and extent-based thresholding. In addition, two left IPL clusters survived only the cluster-extent FWE correction. In contrast to Noonan et al. (2013), we did not find convergent activation in ACC, bilateral SFG, left AG, right IPL/SPL, and left anterior MTG.

The extent to which brain regions engaged in semantic control overlap with those engaged in general semantic cognition (including both representation and control processes), and domain-general executive processes (i.e., the MDN) are illustrated in Figure 3 and is largely the same as that recently highlighted by Jackson (2021).
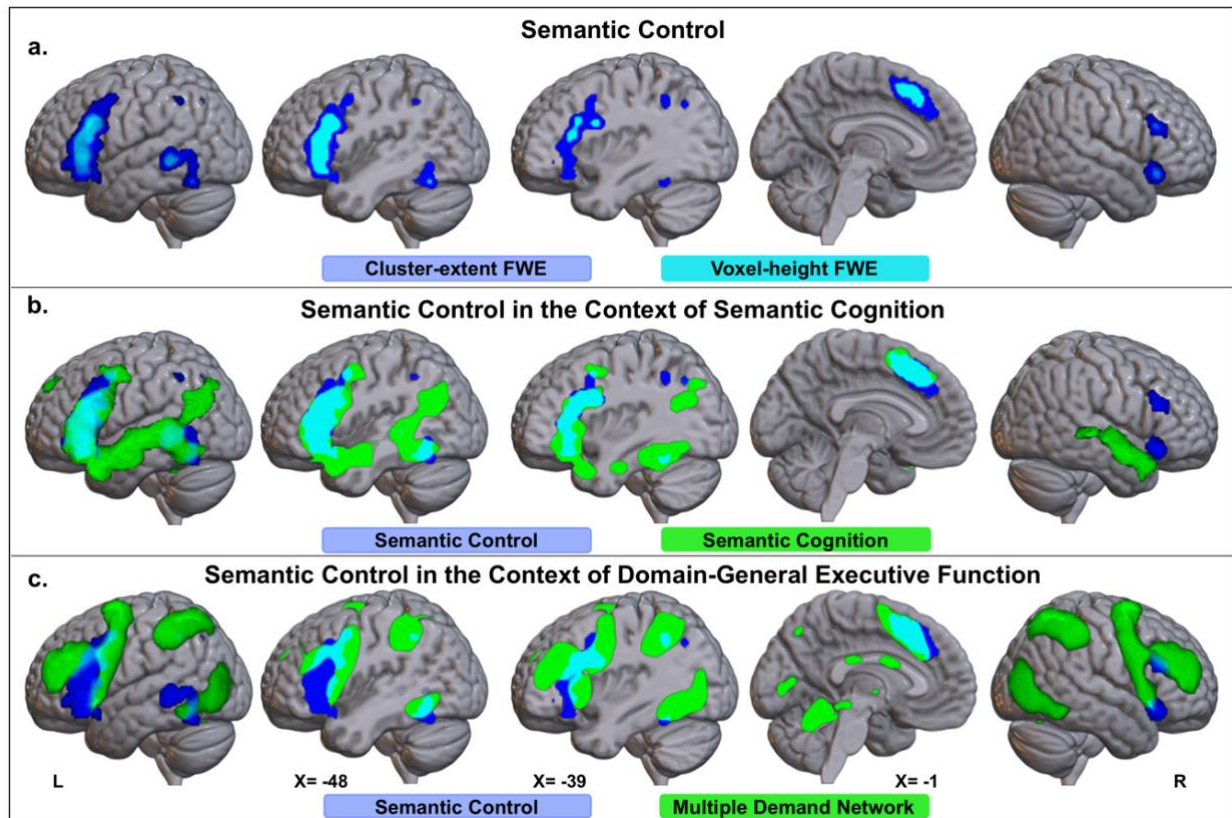
Figure 3. The neural network engaged in semantic control: (a) Binarized ALE maps showing statistically significant convergent activation across 92 experiments contrasting high > low semantic control thresholded using cluster-extent FWE correction of p < .05 with a cluster-forming threshold of p < .001 (blue) and voxel-height FWE correction of p < .05 (cyan). (b) The binarized semantic control map (blue) overlaid with a binarized ALE map of convergent activation across N = 415 semantic > non-semantic contrasts generated in Jackson, 2021 (green); overlap is shown in cyan. (c) The binarized semantic control map (blue) overlaid with a mask of the multiple-demand network (MDN) generated in Federenko et al., 2013 (green) by contrasting hard > easy versions of seven diverse cognitive tasks; overlap is shown in cyan. The lateral views, which show projections on the cortical surface, are accompanied by brain slices at the sagittal midline and also coplanar with the peak of the left STG (X = -48) and left IFG (X = -39) clusters that overlapped across all four social domains (Table S1.5).

3.3. Neural substrates shared by semantic control and social cognition

3.3.1. Theory of mind

Overlap between the neural network underpinning semantic control (i.e., SCN & regions of the MDN) and the ToM network was found in 8 clusters located in the left IFG (including pars orbitalis and triangularis and extending to the precentral gyrus) and, to a smaller extent, the right IFG, the left dorsal precentral gyrus, SMA, left pMTG, left superior temporal pole and the left fusiform gyrus (Figure 4a, Table S3.1.1). The results of the conjunction between semantic control and false belief reasoning can be found in Section S3.1 of SI1 and Table S3.1.2 of SI3. This analysis revealed overlapping activation in the pMTG, but not in the SMA or lateral frontal cortex.

3.3.2. Trait inference

Brain areas involved in both semantic control and trait inference included bilateral IFG (pars orbitalis), SMA and dmPFC (Figure 4b, Table S3.2).

3.3.3. Empathy for emotions

The neural network underpinning semantic control overlapped with the areas engaged in empathy for emotions in bilateral IFG (pars orbitalis and pars triangularis) and SMA (Figure 4c, Table S3.3).

3.3.4. Empathy for pain

Overlapping activation between empathy for pain and semantic control was revealed in left IFG (pars orbitalis and pars triangularis), right IFG (pars orbitalis), left precentral gyrus, bilateral insula, SMA and left posterior ITG (extending towards pMTG) (Figure 4d, Table S3.4).

3.3.5. Moral reasoning

Overlapping activation in response to semantic control and moral reasoning included left insula (extending to pars orbitalis of the IFG), right IFG (pars orbitalis), left IFG (pars opercularis and pars triangularis), the left precentral gyrus and ACC (Figure 4e, Table S3.5).

Overall, the neural network engaged in semantic control overlapped with the neural networks underpinning all four social domains in the bilateral IFG and, in particular, pars orbitalis. Except for moral reasoning, overlapping activation was also found in the SMA. In

the left pMTG, we found a large area of overlap between semantic control and ToM and some evidence of overlap between semantic control and empathic processing.
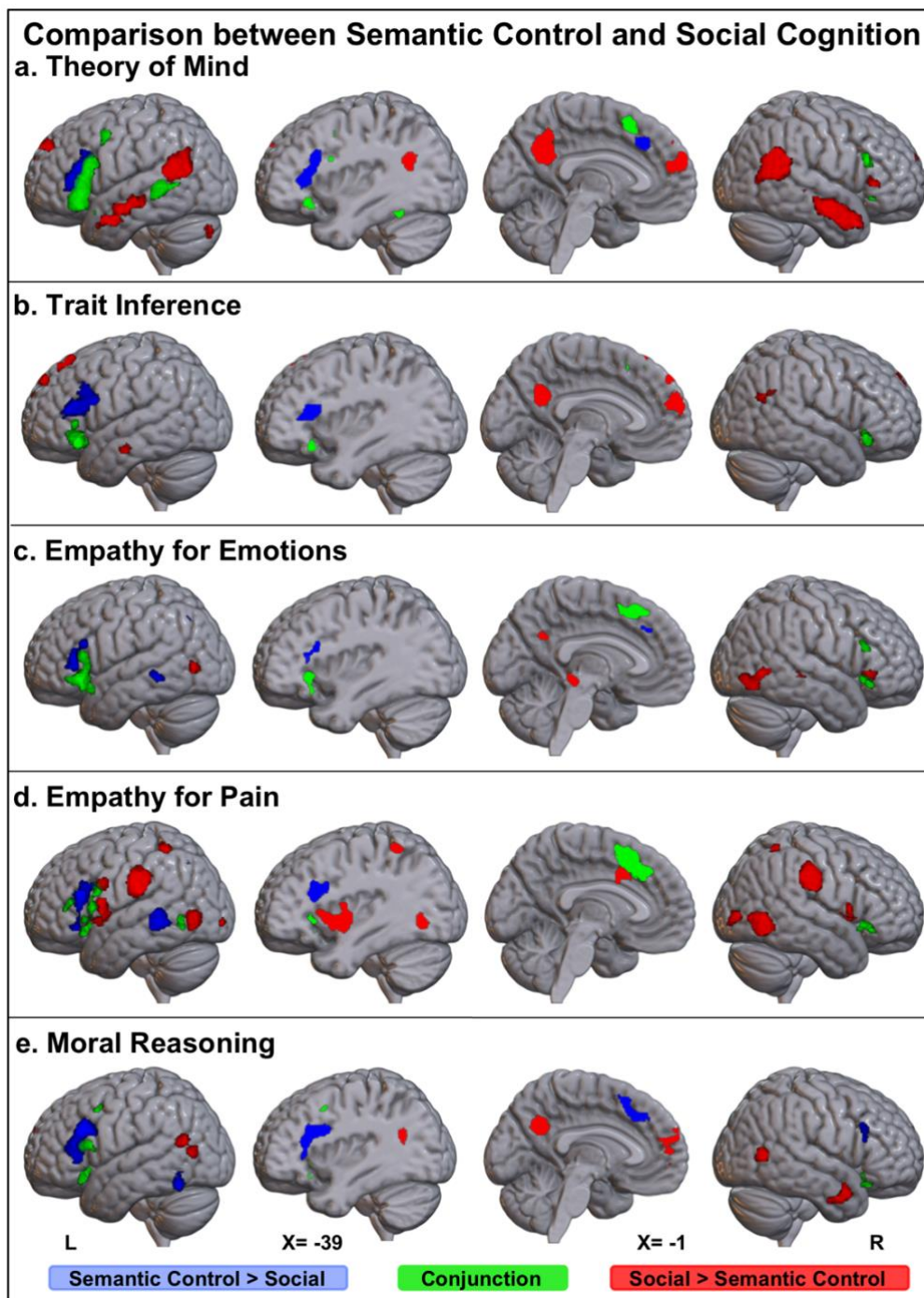
Figure 4. Results of the contrast (blue, red) and conjunction (green) analyses between the ALE maps associated with semantic control and each social domain: a) Theory of Mind b) Trait Inference c) Empathy for Emotions d) Empathy for Pain and e) Moral Reasoning. The contrast maps were thresholded with a cluster-forming threshold of $p < .001$ and a minimum cluster size of 200 mm^3. The lateral views, which show projections on the cortical surface, are accompanied by brain slices at the sagittal midline and also

coplanar with the peak of the left IFG cluster (X = -39) that overlapped across all four social domains (Table S1.5).

3.4. Explicit versus implicit social cognition

Further to the meta-analyses above, we compared activation associated with implicit and explicit paradigms for studying empathy for emotions, empathy for pain and moral reasoning. The results of independent analyses are displayed in Figure 5a–c and Tables S4.1.1–S4.1.6). Conjunctions and formal contrasts are displayed in Figure 5d–f and Tables S4.2.1–S4.2.3). The only notable difference between activation associated with explicit and implicit paradigms, as identified by these formal comparisons, was in the case of empathy, with a small cluster in the dmPFC showing increased convergence for explicit as compared to implicit empathy for pain (see Section S3.4.1. of SI1). In addition, we conducted exploratory cluster analyses to investigate whether the explicit and implicit experiments contributed similarly to each of the significant ALE clusters found for each social domain. In summary, these analyses (Figure S9) revealed that in the case of all social domains, implicit and explicit experiments contributed equally to most clusters (see Section S3.4.2. of SI1 for a more detailed description).
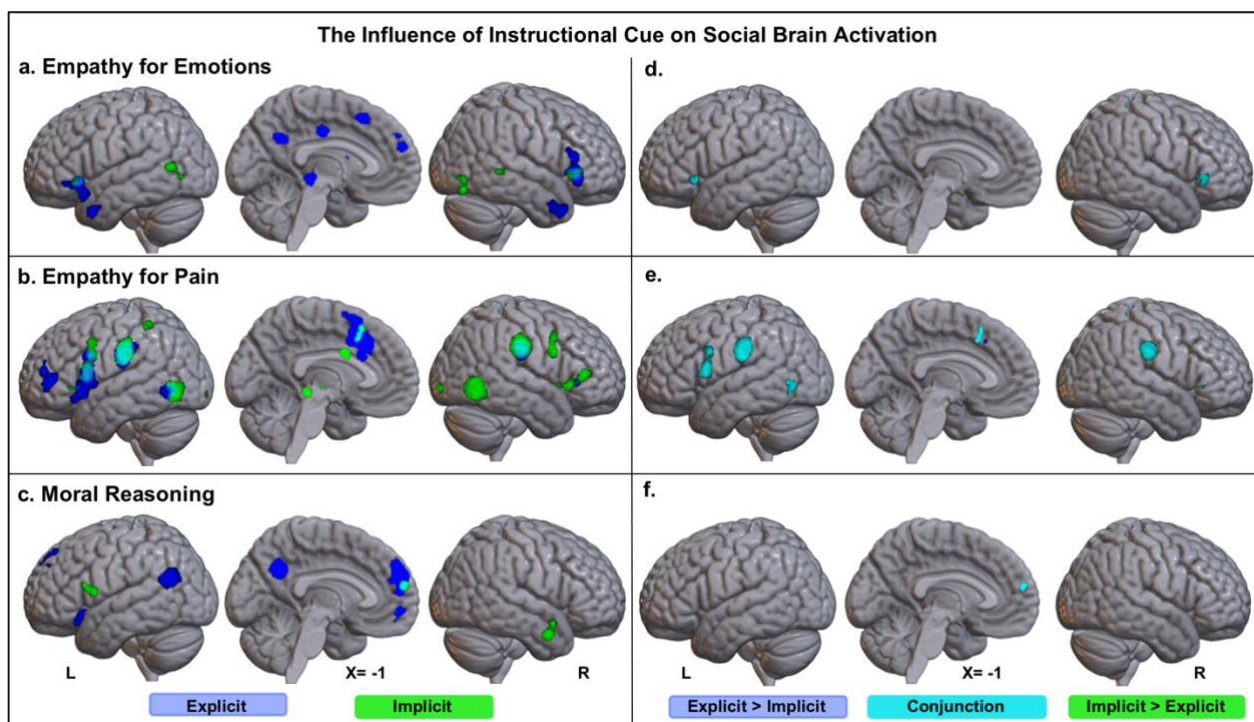
Figure 5. The left panel displays the binary ALE maps showing statistically significant convergent activation resulting from independent meta-analyses on explicit (blue) and implicit (green) studies on a) Empathy for Emotions, b) Empathy for Pain and c) Moral Reasoning. The ALE maps were thresholded using an FWE corrected cluster-extent threshold of $p < .05$ with a cluster-forming threshold of $p < .001$. The right panel displays the results of the contrast (dark blue, green) and conjunction (cyan) analyses between the ALE maps associated with explicit and implicit instructions. The contrast maps were thresholded at $p < .001$ and using a minimum cluster size of 200 mm^3. The lateral views, which show projections on the cortical surface, are accompanied by brain slices at the sagittal midline.

3.5. The relationship between cognitive effort and brain regions engaged during social cognitive tasks

The above-reported conjunction analyses suggest that social cognition engages regions associated with semantic control. In these analyses, we took a pooled approach which involved collapsing over many different comparisons between social and non-social tasks and ignoring subtler differences between experimental and baseline conditions. The key advantage of this approach is that it identifies activation that is generalisable across highly variable experimental conditions. However, ignoring experimental differences precludes a determination of more specific factors driving a given region's involvement. In particular, it is not possible to directly

infer from the above results that semantic control regions are specifically being engaged by the cognitive control demands of social tasks. Therefore, to address this issue, we performed a set of exploratory analyses to determine whether the IFG and pMTG regions are sensitive to the degree of cognitive effort required to complete social tasks. While these analyses cannot disentangle semantic control from other forms of control, they represent a further initial step towards confirming a role of semantic control regions in social regulatory processes. To this end, we took experiments that used explicit paradigms and, on the basis of reported inferential statistics regarding participants' reaction/decision times, categorised them according to whether the experimental condition was more difficult than the control condition (E>C), experimental and control conditions were equally difficult (E=C), or the experimental condition was easier than the control condition (C>E). In the subsequent set of analyses we worked with the premise that in the case of E=C experiments and C>E experiments, activation associated with cognitive effort that is common to both the experimental and control conditions is subtracted away (along with activation specific to the control condition). In contrast, E>C experiments preserve activation associated with cognitive effort that is specific to the experimental condition. Therefore, a contrast analysis pitting E>C experiments against either C>E or E=C experiments will reveal activation associated with cognitive effort specific to the social domain. A conjunction will reveal activation associated with social processing irrespective of task difficulty.

There was only enough information regarding behavioural data to allow for sufficiently powered analyses in the case of ToM (Figure S10 of SI1) where there were 26 E>C ToM experiments and 25 E=C ToM experiments. A conjunction analysis of E>C and E=C experiments yielded common activation in the left IFG (pars orbitalis and pars triangularis), dmPFC, precuneus, bilateral anterior MTG, right pMTG and left SMG (cyan in Figure 6a; Table S5.3) which we interpret as regions engaged in ToM irrespective of task difficulty. Interestingly, a contrast of E>C with E=C ToM experiments revealed differential activation in the left pMTG, an area implicated in semantic control. The full reports of these analyses, including prerequisite independent ALE analyses on the E>C ToM and E=C ToM experiments, can be found in Tables S5.1–S5.4. For completeness, we also analysed C>E ToM experiments, but the sample size (N = 14) was smaller than required to be sufficiently powered (Eickhoff et al., 2016) and therefore the result should be interpreted with caution (Figure 6a, Table S5.4).

Secondly, we conducted exploratory analyses to assess whether E>C, E=C or C>E ToM experiments were equally likely to contribute to each activation cluster (Figure 6b). The clusters were identified in an independent ALE analysis of ToM experiments limited to those

for which the behavioural information was known (Figure 6c; Table S5.5). We expected clusters within brain areas that have a cognitive control function to have a disproportionate contribution from experiments in which the experimental task was more difficult than the control condition. To assess this, we conducted Fisher's exact tests and then interrogated significant main effects through post-hoc pairwise comparisons and using false-discovery-rate adjustments for multiple comparisons. This cluster analysis revealed that E>C, E=C and C>E experiments contributed equally to mPFC ($p = .67$), precuneus ($p = .8$), right anterior MTG ($p = .85$), left pMTG ($p = .74$), right pMTG ($p = .15$) and right IFG ($p = .15$). Contributions to the left IFG cluster depended on the difficulty category ($p < .001$) and pairwise comparisons indicated that the C>E experiments contributed significantly less peaks compared to E>C ($p = .001$) and E>C ($p = .046$) experiments. Contributions to the left anterior MTG cluster also depended on the difficulty category ($p = .043$) and pairwise comparisons indicated that the C>E experiments contributed fewer peaks compared to E>C, but this effect did not survive correction for multiple comparisons ($p = .06$). These results suggest that the left IFG is particularly sensitive to cognitively-challenging ToM processing.
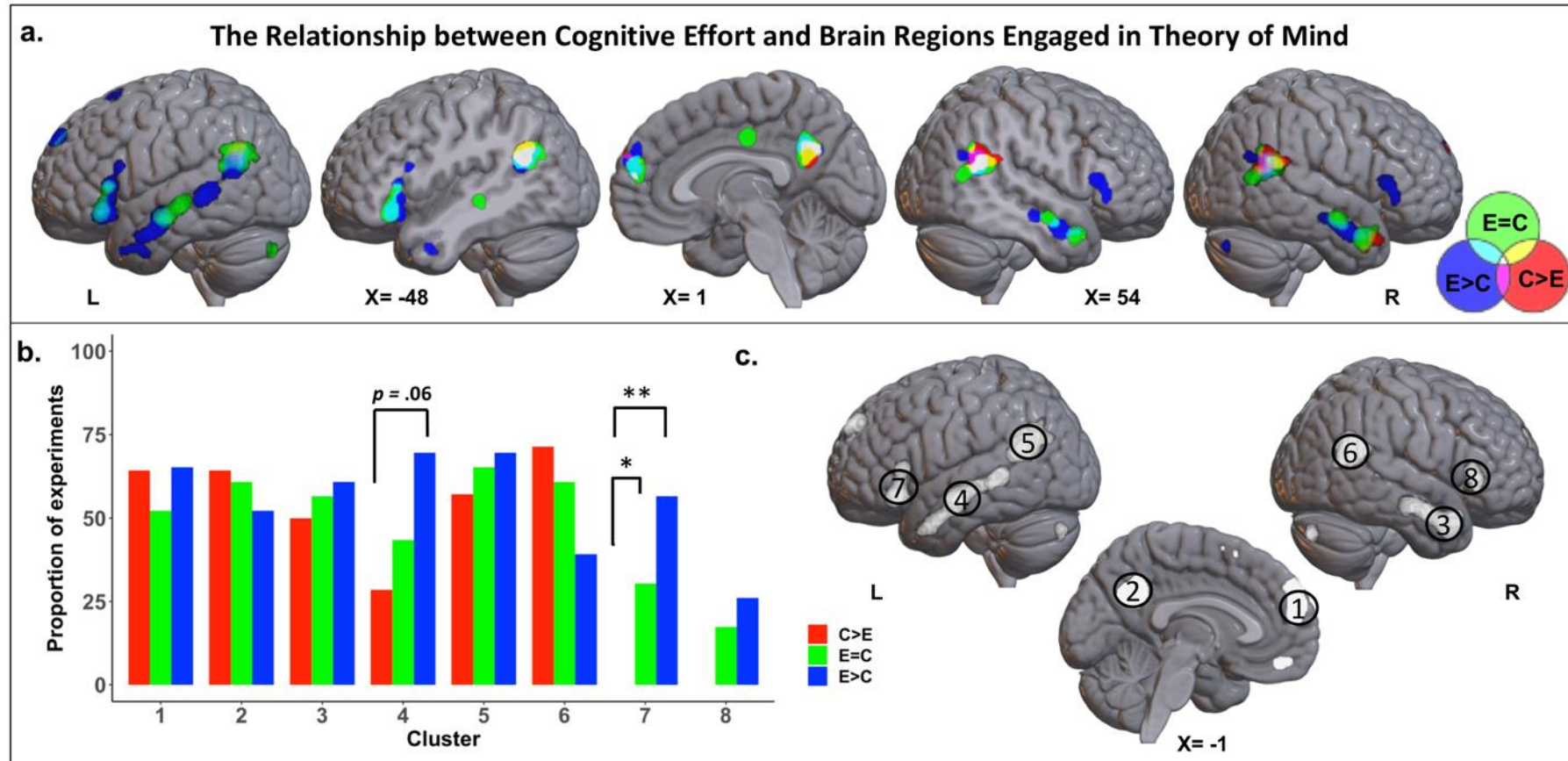
Figure 6. Results of exploratory analyses investigating the effect of task difficulty on ToM activation: (a) Binary ALE maps showing statistically significant convergent activation resulting from independent meta-analyses of three subsets of explicit ToM studies characterized by experimental conditions that were harder than the control task (E>C; N=26; blue), experimental and control conditions that were equally difficult (E=C; N=27; green) and control conditions that were harder than the experimental condition (C>E; N=14; red) as indexed by participant reaction times. The ALE maps were thresholded using an FWE corrected cluster-extent threshold at $p < .05$ with a cluster-forming threshold of $p < .001$. The lateral views, which show projections on the cortical surface, are accompanied by brain slices at the

sagittal midline and also coplanar with the peak of the left IFG cluster (X = -39) that overlapped across all four social domains (Table S1.5) and the right pSTG cluster from the ToM meta-analysis (Table S1.1.1). (b) The results of the cluster analyses where bars represent the proportion of experiments in each difficulty category contributing to clusters of interest resulting from an ALE analysis of N = 60 ToM meta-analysis which included E>C, E=C and C>E experiments. (c) Binary ALE map showing statistically significant convergent activation across ToM experiments limited to those for which the behavioural information was known – this map represented the basis of the cluster analysis. The ALE map was thresholded using an FWE corrected cluster-extent threshold at $p < .05$ with a cluster-forming threshold of $p < .001$; ** $p < .001$ * $p < .05$.

4. Discussion

Although some contemporary theories of social cognition acknowledge the importance of control, or regulatory processes (Adolphs, 2010; Amodio & Cikara, 2021; Frith & Frith, 2012), many key questions remain about their exact nature and neural underpinnings. In the present study, we began to address three such questions: a) whether multiple forms of cognitive control contribute to social cognition, b) whether these control processes are ubiquitously involved in or selectively engaged by certain social abilities, and c) whether this engagement is dependent on specific task demands (e.g., instructional cue) (Binney & Ramsey, 2020). Specifically, we set out to investigate whether brain regions implicated in the controlled retrieval and selection of conceptual knowledge - particularly the IFG and pMTG comprising the SCN (Jefferies, 2013; Lambon Ralph et al., 2017) - contribute to social processing. We simultaneously applied this question to multiple sub-domains of social cognition so that we could assess the extent to which involvement is general, or specific to certain types of social tasks and/or abilities. And we adopted a formal meta-analytic approach to extracting reliable trends from across a large number of functional neuroimaging studies and overcome the limitations of individual experiments (Cumming, 2014; Eickhoff et al., 2012). We found that theory of mind, trait inference, empathy, and moral reasoning commonly engage a core social network that includes the left IFG, precuneus and, when more liberal thresholds are applied, the right IFG, mPFC, bilateral ATL and left pMTG/AG. Moreover, the IFG (particularly the pars orbitalis) region greatly overlapped with that implicated in an independent meta-analysis of neuroimaging studies of semantic control. Further, exploratory analyses suggest that both the left anterior IFG and the left posterior MTG (at a position just anterior to the 'temporoparietal junction')

are sensitive to executive demands of social tasks. We interpret our overall findings as supportive of the hypothesis that the SCN supports social cognition via a process of controlled retrieval of conceptual knowledge. This aligns with a broader proposal in which social cognition is described as a flavour of domain-general semantic cognition and relies on the same basic cognitive and brain systems (Binney & Ramsey, 2020).

4.1. Cognitive control in social cognition

4.1.1. The contribution of semantic control

Alternative theories and existing findings regarding cognitive control in social cognition point to distinguishable mechanisms for monitoring conflict and errors, and for implementing or inhibiting responses (see, for example, Amodio, 2014). However, the exact nature of the information or processes being controlled is not clear. Involvement of the SCN in social cognition suggests that it is, at least in part, related to a controlled attribution of meaning to stimuli and experiences, and to the production of task-appropriate meaning-imbued behaviour (Corbett et al., 2015; Lambon Ralph et al., 2017). Within the broader literature regarding semantic control, a key distinction has been drawn between a) top-down goal-directed retrieval and b) post-retrieval selection of goal-relevant semantic knowledge (Badre et al., 2005; Jefferies, 2013; Thompson-Schill et al., 1997), and it has been suggested that both of these two control mechanisms contribute significantly to interpersonal interactions (Binney & Ramsey, 2020; Satpute et al., 2014). Studies of semantic cognition suggest that 'selection' is engaged when bottom-up, automatic activation of conceptual knowledge results in multiple competing semantic representations and/or responses. Social interactions frequently involve subtle or ambiguous cues, such as neutral facial expressions and bodily gestures, and/or conflicting cues (e.g., sarcasm). This causes semantic competition that can only be resolved by taking into account the wider situational and linguistic context and/or prior knowledge about the speaker (Aviezer et al., 2008; Pexman, 2008). Controlled retrieval processes, on the other hand, are engaged when automatic semantic retrieval fails to activate the semantic information necessary for the task at hand. This may occur frequently in social interactions, and particularly with less familiar persons, because of a preponderance of surface features (e.g., physical characteristics) over less salient features (e.g., personality traits, preferences, and mental states). To avoid exchanges that are deemed superficial at best, controlled retrieval must be used to bring to the fore person-specific but also context-relevant semantic information. On the basis of observations in other domains, it is possible to make some predictions about what social behaviour might look like when these semantic regulatory processes fail. For example,

semantic errors in language production (as opposed to phonological errors, for example) arise because of demanding (e.g., speeded) testing conditions, as well as brain pathology (Hodgson & Lambon Ralph, 2008; Jefferies & Lambon Ralph, 2006), and can be categorised according to three types: superordinate (e.g., saying "animal" in response to a picture of a horse), coordinate (e.g., saying the name of an incorrect but related, often more familiar concept, e.g., "cat") and associative errors (e.g., "stable"). These errors reflect a failure to access the correct meaning associated with a stimulus. When the ability to impose semantic control is compromised during social interactions one might observe similar types of errors; that is behavioural responses that are incongruous with, albeit distantly semantically related to incoming interpersonal signals. Further, one would predict that these errors are less likely when contextual anchors constrain the possible meanings and reduce the reliance on semantic control. In line with this, a recent study has demonstrated effects of impaired semantic control on emotion perception (Souter et al., 2021).

There is now over a decade's worth of multi-method evidence that semantic control is underpinned by the left IFG and the left pMTG (Jefferies, 2013; Lambon Ralph et al., 2017). Research is now aimed at understanding the neural mechanisms by which these regions modulate semantic processing. One recent proposal is that it involves coordination of spreading activation across the semantic representational system (Chiou et al., 2018). According to the hub-and-spoke theory of semantic representation (Lambon Ralph et al., 2017), coherent concepts are represented conjointly by a central supramodal semantic 'hub' located in the ATLs, as well as multiple distributed areas of association cortex ('spokes') that represent modality-specific information (e.g., visual features, auditory features, verbal labels, etc.). Chiou et al. (2018) showed that the left IFG could be imposing cognitive control by flexibly changing its effective connectivity with the hub and spoke regions according to task characteristics; the IFG displayed enhanced functional connectivity with the 'spoke' region that processes the most task-relevant information modality. A similar proposal has been made for the contribution of domain-general cognitive control systems to social information processing. Zaki et al. (2010) found that, in the presence of conflicting social cues, right IFG activity becomes functionally coupled with the brain areas associated with processing the particular cue type the participant chose to rely on to make inferences about emotional states. On this basis, they proposed that cognitive control areas upregulate activation within systems that represent social cues that are currently most relevant to the task. Consistent with this, a further study found evidence to suggest that the left IFG downregulates neural activation

associated with task-irrelevant self-referential information when the task required reference to others (and vice versa) (Soch et al., 2017).

An important feature of semantic processing is the ability to accommodate new information that emerges over extended periods of time and update our internal representation of the current 'state of affairs' in the external world according to contextual changes. This is particularly important for navigating social dynamics which are liable to abrupt and sometimes extreme changes in tone. For instance, imagine being in a bar and having your attention drawn to someone standing suddenly and picking up a glass. One might reasonably infer that this person is thirsty. That is until they proceed to walk towards a group of noisy sports fans rather than the bartender. In this case, you will likely adapt your interpretation and engage in a pre-emptive defensive stance. Recent research suggests that this ability to update depends, at least in part, on the IFG, as well as the mPFC and ventral IPL (also see Section 4.2.2) (Branzi et al., 2020). Likewise, Lavoie et al. (2016) showed that, during a ToM task, activation of the left IFG and pMTG is associated with contextual adjustments of mental state inferences (and also more general physical inferences) although not the representation of mental states specifically. Left IFG activation has also been observed when newly-presented information requires one to update the initial impression formed of another person (e.g., Mende-Siedlecki, Baron, et al., 2013; Mende-Siedlecki, Cai, et al., 2013; Mende-Siedlecki & Todorov, 2016).

4.1.2. The wider contribution of executive processes

According to Lambon Ralph, Jefferies, and colleagues, the executive component of semantic cognition comprises both semantic control and other domain-general processes (Lambon Ralph et al., 2017). The latter includes top-down attentional control and working memory systems that support goal-driven behaviour irrespective of the task domain (i.e., perceptual, motor or semantic). These processes recruit nodes of the MDN (Duncan, 2010), which include the precentral gyrus, MFG, IPS, insular cortex, pre-SMA and adjacent cingulate cortex (Assem et al., 2020; Fedorenko et al., 2013). In terms of organisation, the SCN appears to be nested amongst domain-general executive systems (Wang et al., 2020) and could play a role in mediating interactions between the MDN and the semantic representational system (Davey et al., 2016; Lambon Ralph et al., 2017). In line with this general perspective, we expected MDN regions to be reliably engaged by all four social sub-domains explored in the present meta-analyses. While there was evidence of engagement of the MFG, the pre-SMA, ACC, insula and IPS in some of the social sub-domains, MDN regions were not part of the core social processing network identified by the overlay conjunction analysis. This could reflect the fact

that the majority of contrasts included in our meta-analyses employed high-level control conditions that were well-matched to the experimental conditions in terms of general task requirements, and thus, most activation associated with general cognitive demands had been subtracted away. Consistent with this notion is the fact that studies contrasting social tasks with lower-level control conditions (e.g., passive fixation) find more extensive MDN activation in ToM (Mason et al., 2008; Mier et al., 2010), trait inference (Chen et al., 2010; Hall et al., 2012), empathy (De Greck et al., 2012; Tamm et al., 2017) and moral reasoning (Reniers et al., 2012). The role of the MDN in social cognition is otherwise becoming well-established, and it has been found to be sensitive to difficulty manipulations in social tasks, showing increased activation in response to conflicting social cues (Cassidy & Gutchess, 2015; Mitchell, 2013), social stimuli that violate expectations (Cloutier et al., 2011; Hehman et al., 2014; Ma et al., 2012; Weissman et al., 2008) and increasing social working memory load (Meyer et al., 2012).

Finally, it is important to note that, although both MDN and the SCN co-activate in social and semantic tasks, the nature of their specific contributions *and* their anatomy are at least partially dissociable. The MDN is associated with the implementation of top-down constraints to facilitate goal-driven aspects of processing that is not limited to the semantic domain (Duncan, 2013; Fedorenko et al., 2013; Whitney et al., 2012). In contrast, the engagement of the anterior ventrolateral IFG (pars orbitalis) and the left pMTG appear specific to the semantic domain and, in particular, controlled semantic retrieval (Badre & Wagner, 2007; Dobbins & Wagner, 2005; Whitney et al., 2012). Unlike the MDN, they do not appear to respond to challenging non-semantic tasks (Gao et al., 2021; Hodgson et al., 2021; Noonan et al., 2013; Whitney et al., 2012). Further, tasks associated with low conceptual retrieval demands but a requirement for response inhibition engage the MDN but do not engage the SCN, even if conceptual knowledge is used to guide responses (Gonzalez Alam et al., 2018).

4.1.3. Double-route vs single-route cognitive architecture of social cognition

A secondary aim of the present study was to address a pervasive distinction in the social neuroscientific literature between automatic and controlled processes (Adolphs, 2010; Happé et al., 2017; Lieberman, 2007). Some authors have argued that automatic and controlled social processes are mutually exclusive of one another and draw upon distinct cortical networks (Forbes & Grafman, 2013; Lieberman, 2007). The alternative is a single-route architecture where the degree to which behaviours have particular attributes (e.g., speed, effort, intentionality) does not reflect the involvement of one system and not another, but quantitative differences in the extent to which the control system interacts with the representational system

in order to produce context-/task- appropriate responses (Binney & Ramsey, 2020). Our results are consistent with the latter perspective. The brain regions reliably activated in response to explicit instructions and those associated with implicit instructions revealed more overlap than discrepancy across empathy and moral reasoning tasks. Notably, this overlap included brain areas associated with executive functions: the bilateral IFG in the case of empathy for emotions and bilateral IFG and dmPFC in the case of empathy for pain. Moreover, cluster analyses of the ALE maps associated with the four social domains suggest that studies using explicit and implicit paradigms (which are assumed to engage controlled and automatic processing respectively) contributed equally to most activation clusters, including those in brain regions associated with control processes. Contrary to the predictions of dual-process models, these findings suggest that common neural networks contribute to both explicit and implicit social processing (Van Overwalle & Vandekerckhove, 2013). Furthermore, exploratory analyses suggest that both the left anterior IFG and the pMTG are sensitive to executive demands of social tasks. Overall, we argue that these results support the existence of a single-route cognitive architecture wherein the contribution made by control mechanisms to implicit and explicit social processing reflects cognitive effort demanded by the task at hand. This follows similar proposals put forth specifically in the domain of ToM (Carruthers, 2016, 2017).

4.2.   Beyond cognitive control

Our findings converged upon four further regions that have been strongly linked with key roles in social cognition: the mPFC (including the anterior cingulate), the precuneus, the 'temporoparietal junction' (TPJ), and the ATL. We briefly discuss the putative role of each of these regions below.

4.2.1.   The 'temporo-parietal junction'

A region often referred to as the 'temporo-parietal junction' (TPJ) has been subject to an elevated status within the social neurosciences. In particular, the right TPJ has been attributed with a key role in representing the mental states of others (Saxe & Wexler, 2005). In line with previous meta-analyses (Bzdok et al., 2012; Molenberghs et al., 2016; Schurz et al., 2013, 2014, 2020), our results reveal a bilateral TPJ region that is reliably involved in ToM tasks. In the left hemisphere, an overlapping area is also implicated in trait inference, moral reasoning and, when a more lenient threshold was applied, empathy for emotions, which is suggestive of a broader role of the left TPJ in social cognition. In contrast, the right TPJ showed more limited

overlap, being reliably engaged only by ToM and trait inference tasks, which is suggestive of a more selective role of the right TPJ in social cognition.

The TPJ encompasses a large area of cortex that is poorly defined anatomically and seems to include parts of the AG, SMG, STG and MTG (Schurz et al., 2017). Moreover, this area is functionally heterogeneous and has been associated with a variety of cognitive domains including but not limited to attention, language, numerosity, episodic memory, semantic cognition and social perception (Binder et al., 2009; Decety & Lamm, 2007; Deen et al., 2015; Humphreys & Lambon Ralph, 2015; Igelström & Graziano, 2017; Özdem et al., 2017; Quadflieg & Koldewyn, 2017). While there is some indication that the function of the TPJ may be dependent on the hemisphere (Numssen et al., 2021), many cognitive domains, including ToM, are associated with bilateral TPJ activation. Our results at least seem to suggest dissociable roles of pMTG and a more posterior TPJ region; while the left pMTG is activated within both semantic control and ToM studies, a separate and more posterior STG (TPJ) area located closer to SMG/AG was reliably engaged by three of the social domains, but not studies of semantic control. Furthermore, the results suggest that the left pMTG is sensitive to the difficulty of ToM tasks while the bilateral pSTG (TPJ) region is not.

This finding is generally in line with previous research suggesting a functional dissociation between the left pMTG and the left ventral IPL/AG regions. From one perspective, the activation of both regions appears to be positively associated with semantic tasks (Binder et al., 2009). However, the left pMTG shows increased activation to difficult relative to easier semantic tasks (Jackson, 2021; Noonan et al., 2013), unlike the ventral IPL/AG which has been shown to deactivate to semantic tasks when they are contrasted against passive/resting conditions where there may be greater opportunity for spontaneous semantic processing or 'mind-wandering' (Humphreys et al., 2015; Humphreys & Lambon Ralph, 2015). Moreover, Davey et al. (2015) found that TMS applied to pMTG disrupted processing of weak semantic associations more than for strong associations, whereas TMS applied to AG had the opposite effect. Based on these and similar observations it has been suggested that the ventral IPL/AG has a role in the automatic retrieval of semantic information.

4.2.2. The default mode network

The pSTG/AG and the mPFC and precuneus regions we identified as part of the core social cognition network are also considered part of the default-mode network (DMN) (Buckner et al., 2008; Spreng & Andrews-Hanna, 2015). The DMN is a resting-state network, meaning that it is a group of regions consistently co-activated without the requirement of an explicit task. It

is proposed that it is ideally suited for supporting self-generated internally-oriented as opposed to externally-oriented cognition (i.e., it is decoupled from sensory processing; Margulies et al., 2016; Smallwood et al., 2013). Some of these regions (e.g., the AG and mPFC) have also been implicated in processes that allow the integration of information over time (Huey et al., 2006; Humphreys et al., 2021; Ramanan et al., 2018; Ramanan & Bellana, 2019). These purported functions are all presumably important for social and more general semantic processing (see Section 4.1.1.) and likely involve domain-general mechanisms (also see Van Overwalle, 2009). However, the degree to which regions implicated in the DMN and those implicated in social and/or semantic cognition do or do not overlap is contentious and much is left to be gleaned regarding the relationship between these systems (Jackson et al., 2019, 2021a; Mars et al., 2012).

4.2.3. The anterior temporal lobe

Our findings implicate the lateral anterior temporal lobe (ATL), and particularly the dorsolateral STG/temporal pole (BA 38) and middle anterior MTG/STS, in all the socio-cognitive domains investigated, except for empathy for pain. Exploratory cluster analyses revealed that ATL engagement is not dependent on instructional cue or task difficulty, and thus it appears to serve a role other than control.

A key contribution of the ATL to social-affective behaviour has been recognised by comparative and behavioural neurologists for well over a century, owed at first to the acclaimed work of Brown and Schafer (1888) and, later, Klüver and Bucy (1939) who provide detailed reports of profound social and affective disturbances in non-human primates following a bilateral, full depth ATL resection. These observations are mirrored in descriptions of neurogenerative patients that associate progressive ATL damage with a wide range of socio-affective deficits (Binney, Henry, et al., 2016; Chan et al., 2009; Ding et al., 2020; Perry et al., 2001), including impaired emotion recognition (Lindquist et al., 2014; Rosen et al., 2004) and empathy (Rankin et al., 2005), impaired capacity for ToM (Duval et al., 2012; Irish et al., 2014), and a loss of person-specific knowledge (Gefen et al., 2013; Snowden et al., 2004, 2012). Over the past 10 years, there been a growing acceptance of the central role played by the ATL within the social neurosciences (Olson et al., 2013) and it now features prominently in some neurobiological models of face processing (Collins & Olson, 2014), ToM (Frith & Frith, 2006), moral cognition (Moll et al., 2005), and emotion processing (Lindquist et al., 2012). It has also been pinpointed as a key source of top-down influence on social perception (Freeman & Johnson, 2016). One influential account of social ATL function proposes a

domain-specific role in the representation of social knowledge, including person knowledge, and other more abstract social concepts (Olson et al., 2013; Thompson et al., 2003; Zahn et al., 2007).

A parallel line of research focused upon general semantic cognition has given rise to an alternative, more domain-general account of ATL function; there is a large body of convergent multi-method evidence from patient and neurotypical populations in support of a role of the ATL in the formation and storage of all manner of conceptual-level knowledge (Lambon Ralph et al., 2017). Research efforts have therefore recently begun to ask whether the purported roles of the ATL in both social and semantic processes can be reconciled under a single unifying framework (Binney, Hoffman, et al., 2016; Rice et al., 2018). Some clues already exist in the aforementioned work of Klüver and Bucy (1939), who observed a broader symptom complex comprising multimodal semantic impairments, including visual and auditory associative agnosias, that might explain rather than just co-present with social-affective disturbances. More recent work that leverages the higher spatial resolution of functional neuroimaging in humans has revealed a ventrolateral ATL region that responds equally to all types of concepts, including social, object and abstract concepts, be they referenced by verbal and/or non-verbal stimuli (Binney, Hoffman, et al., 2016; Rice et al., 2018; Zahn et al., 2007). Activation of the dorsal-polar ATL, on the other hand, appears to be more sensitive to socially-relevant semantic stimuli (Binney, Hoffman, et al., 2016; Rice et al., 2018; Zahn et al., 2007). These observations support a proposal in which the broadly-defined ATL region can be characterised as a domain-general supramodal semantic hub with graded differences in relative specialisation towards certain types of conceptual information (Binney et al., 2012; Binney, Hoffman, et al., 2016; Lambon Ralph et al., 2017; Plaut, 2002; Rice et al., 2015). Our results reveal that the temporal poles are reliably activated across four social domains – moral reasoning, empathy for emotions, ToM and trait inference. They do not, however, provide support for the involvement of the ventrolateral ATL. We argue this is likely due to technical and methodological limitations of the fMRI studies included in the meta-analyses (Visser et al., 2010). Most notably, this includes vulnerability to susceptibility artefacts that cause BOLD signal drop-out and geometric distortions around certain brain areas, including the ventral ATLs (Jezzard & Clare, 1999; Ojemann et al., 1997). Studies that have used PET, which is not vulnerable to such artefacts, or techniques devised to overcome limitations of conventional fMRI (Devlin et al., 2000; Embleton et al., 2010), reveal activation in both the temporal poles and the ventral ATL in response to social stimuli (Balgova et al., 2021; Binney, Hoffman, et al., 2016; Castelli et al., 2002).

4.3. Limitations

Because semantic control demands were not explicitly manipulated in the social contrasts we included, our results cannot directly confirm our hypothesis regarding the specific contribution made by the SCN in social cognition. Our conclusions rely on an assumption that overlap reflects a generalised neurocomputation upon which both semantic control and social processing rely. The alternative explanation is that overlapping activation reflects tightly yet separately packed cognitive functions which may only dissociate when investigated at an increased spatial resolution (Henson, 2006; Humphreys et al., 2021). Moreover, we chose to pool across heterogeneous samples of studies to investigate the cognitive domains of interest. The advantage of this approach is that it identifies activation that is generalisable across highly variable experimental conditions and washes out spurious findings associated with idiosyncratic properties of stimuli and/or paradigms. However, the preponderance of specific experimental procedures in each literature addressed still unintentionally led to systematic differences in the characteristics of the studies used to define the different cognitive domains. For example, the semantic control dataset included studies that employed verbal stimuli almost exclusively, while the majority of empathy studies employed non-verbal stimuli. Some of the differences between the associated networks (e.g., in lateralization) might therefore be attributable to verbal processing demands. As is the case with all meta-analyses, therefore, some aspects of our results should be treated with caution.

Another limitation of this study is that most of the experiments included used control conditions that were highly matched to their experimental conditions in terms of the demand for domain-general processes such as cognitive control and semantic processing, and therefore they may have subtracted away much of the activation we were aiming to explore. Despite this, we did find consistent activation of the SCN, particularly the left IFG, across all four social domains. This may be because, although a considerable subset of included experiments had high-matching control conditions, not all may have properly controlled for semantic control demands specifically. An alternative explanation is that processing socially-relevant conceptual knowledge may impose greater demands on the SCN. Consistent with this, it has been shown that processing social concepts relative to non-social concepts led to increased activation of the SCN even when controlling for potentially confounding psycholinguistic factors (Binney, Hoffman, et al., 2016).

4.4. Concluding remarks and future directions

Regions of the SCN are engaged by several types of complex social tasks, including ToM, empathy, trait inference and moral reasoning. This finding sheds light on the nature and neural correlates of the cognitive control mechanisms which contribute to the regulation of social cognition and specifically implicates processes involved in the goal-directed retrieval of conceptual knowledge. Importantly, our current findings and our broader set of hypotheses can be generalised to multiple social phenomena, thereby contributing a unified account of social cognition. Future research will need to establish a causal relationship between the SCN and successful regulation of social processing. This could be done by investigating the capacity for neurostimulation of SCN regions to disrupt social task performance. Similarly, whether SCN regions are sensitive to manipulations of semantic control demands within a social task could be probed directly.

Elucidating the neural bases of social control and representation may help us understand the precise nature of social impairments resulting from damage to different neural systems. For example, our framework (Binney & Ramsey, 2020) predicts that damage to representational areas such as the ATL will impair social information processing irrespective of task difficulty or the need to integrate context. In contrast, we expect that damage to control areas would lead to impaired social processing specifically when it requires selecting from amongst alternative interpretations of social cues, and/or retrieving social information that is only weakly associated with a person or a situation. Damage to perisylvian frontal and/or temporo-parietal areas (comprising the SCN) leads to semantic aphasia, a disorder characterized by impaired access and use of conceptual knowledge (Corbett et al., 2009; Jefferies et al., 2007, 2008; Jefferies & Lambon Ralph, 2006; Noonan et al., 2010). This contrasts with ATL damage which leads to semantic dementia, a condition associated with a loss or degradation of semantic knowledge (including social knowledge; Hodges & Patterson, 2007; Lambon Ralph et al., 2010; Lambon Ralph & Patterson, 2008; Patterson et al., 2007; Rogers et al., 2004). As far as we are aware, the extent to which brain damage that leads to semantic aphasia also affects social abilities has only been formally investigated in the case of emotion recognition (Souter et al., 2021). Some other insights can be found in neurodegenerative patients with prominent frontal lobe damage, where social impairments can be linked to deficits in executive function (Healey & Grossman, 2018; Kamminga et al., 2015). More generally, it will be interesting to discover whether a distinction between knowledge representation and cognitive control can inform our understanding of the precise nature of atypical or disordered social cognition in, for

example, the context of dementia, acquired brain injury, autism spectrum conditions and schizophrenia.

CHAPTER 5

Graded functional organisation in the left inferior frontal gyrus: evidence from task-free and task-based functional connectivity

Parts of this chapter have been prepared for publication as:

Diveica, V., Riedel, M.C., Salo, T., Laird, A.R., Jackson, R.L., & Binney, R.J. (in prep). Graded functional organisation in the left inferior frontal gyrus: evidence from task-free and task-based functional connectivity.

Authors' contributions and acknowledgements:

Veronica Diveica, Dr Richard Binney and Dr Rebecca Jackson conceptualized this study, and interpreted the results. All authors contributed to the development of the methodological approach. Taylor Salo developed the NiMARE software used for meta-analytic analyses. The statistical analyses were performed by Veronica Diveica with the support of Dr Michael Riedel, Dr Taylor Salo, Aaron Owen and computational resources provided by the Supercomputing Wales project. Prof Angela Laird and Julio A. Peraza compiled and provided access to annotations of Neurosynth terms. Veronica Diveica performed the data visualization and wrote the original draft. Dr Richard Binney and Dr Rebecca Jackson edited drafts of the manuscript. Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Open science practices and data/code availability statement

We used open data available via the Human Connectome Project (https://www.humanconnectome.org/) and the NeuroQuery database (https://neuroquery.org/). The code used for data pre-processing and analysis can be accessed at: github.com/DiveicaV/LIFG_Gradients. The gradient and functional connectivity maps can be accessed and visualized via NeuroVault at: https://neurovault.org/collections/ETPEFCDV/. Due to its exploratory nature, the study was not pre-registered.

Abstract

The left inferior frontal gyrus (LIFG) has been ascribed key roles in numerous cognitive domains, including language, executive function and social cognition. However, its functional organisation, and how the specific areas implicated in these cognitive domains relate to each other, is unclear. Possibilities include that the LIFG underpins a domain-general function or, alternatively, that it is characterized by functional differentiation, which might occur in either a discrete or a graded pattern. The aim of the present study was to explore the topographical organisation of the LIFG using a bimodal data-driven approach. To this end, we extracted functional connectivity (FC) gradients from 1) the resting-state fMRI timeseries of 150 participants, and 2) patterns of co-activation derived meta-analytically from task data across a diverse set of cognitive domains. We then sought to characterize the FC differences driving these gradients with seed-based resting-state FC and meta-analytic co-activation modelling analyses. Both data types converged on a FC profile that shifted in a graded fashion along two main organisational axes. An anterior-posterior gradient shifted from being preferentially associated with high-level control networks (anterior LIFG) to being more tightly coupled with perceptually-driven networks (posterior). A second dorsal-ventral axis was characterized by higher connectivity with domain-general control networks on one hand (dorsal LIFG), and with the semantic network, on the other (ventral). These results provide novel insights into a graded functional organisation of the LIFG underpinning both task-free and task-constrained mental states, and suggest that the LIFG is an interface between distinct large-scale functional networks.

1. Introduction

The left inferior frontal gyrus (LIFG) is ascribed a key role in numerous cognitive domains, including language (Friederici, 2011), semantics (Lambon Ralph et al., 2017), action (Papitto et al., 2020), social cognition (Diveica et al., 2021), and executive function (Fedorenko et al., 2013). The extent of this overlap is remarkable; however, what exactly is driving it is unknown. One possibility is that LIFG subserves a singular function which manifests as common activation across domains. Alternatively, detailed exploration of its organisation could reveal subregions with multiple functional specialisations.

Some clues are gleaned from detailed studies of cellular micro-structure and white-matter connectivity that date back to Brodmann and his contemporaries (Bailey & Von Bonin, 1951; Brodmann, 1909). Cytoarchitecture and 'fibrillo-architecture' are proposed to determine a region's functional characteristics by constraining local processing capabilities and the incoming/outgoing flow of information, respectively (Cloutman & Lambon Ralph, 2012; Passingham et al., 2002). Indeed, these data reveal that the LIFG is far from uniform and, instead, comprises at least three sub-regions with distinct cytoarchitecture (Amunts et al., 1999; Schenker et al., 2008; Wojtasik et al., 2020), neurotransmitter receptor distributions (Amunts et al., 2010), and structural connectivity (Anwander et al., 2007; Klein et al., 2007; Neubert et al., 2014; Wang et al., 2020). However, it has thus far proven difficult to map these structural distinctions onto functional topographies derived from neuroimaging data.

Various functional dissociations have been identified within the LIFG by means of functional neuroimaging, including distinctions between semantic and phonological language processes (Devlin et al., 2003), and between memory retrieval and post-retrieval selection (Badre & Wagner, 2007). However, they have arisen primarily from experimental cognitive approaches and limited neuroimaging datasets which are poorly suited to generating unifying accounts that explain multiple phenomena. A promising alternative is to take a large-scale data-driven approach that spans cognitive domains (Genon et al., 2018). On this basis, one might encapsulate the full functional repertoire of a brain region.

Functional connectivity (FC) patterns derived from neuroimaging data could prove useful because they capture the extent to which regional activation covaries over time and, therefore, are sensitive to context-dependent inter-regional interactions. Moreover, they can reveal aspects of the connectome that might not manifest within other modalities; FC can arise between anatomically

remote brain areas without direct structural connections (Damoiseaux & Greicius, 2009; Suárez et al., 2020). The small number of studies that have attempted to divide the LIFG into sub-regions based on FC reveal a heterogenous functional architecture (Clos et al., 2013; Kelly et al., 2010). However, these prior investigations have implemented 'hard' clustering algorithms (Eickhoff et al., 2015), which assume that sharp borders separate intrinsically homogeneous neural regions. This means they may fail to identify graded transitions that (i) could give rise to functionally intermediate areas (Bailey & Von Bonin, 1951; Rosa & Tweedale, 2005) and (ii) have been observed in the connectivity patterns of other brain regions (e.g., Bajada et al., 2017; Cerliani et al., 2012; Jackson et al., 2018, 2020; Tian & Zalesky, 2018), as well as within the cytoarchitecture of transmodal cortex (Brodmann, 1909). Therefore, the possibility of graded functional differences in the LIFG remains unexplored.

Insights into the nature of spatial transitions in cortical organisation, or *gradients*, can be gleaned using an emergent analytical approach (Bajada et al., 2020; Huntenburg et al., 2018). A key feature of gradient analyses is that they do not presuppose the nature of variation and, therefore, can be used to demonstrate both graded changes and discrete boundaries (Bajada et al., 2017; Jackson et al., 2018; Johansen-Berg et al., 2004). Moreover, they can distinguish between superimposed but orthogonal spatial dimensions of functional variation which might otherwise appear as a singular aspect of organisation (Haak et al., 2018). Despite these advantages, the gradient approach has not yet been applied to studying the connectivity of the LIFG.

Therefore, our aim was to use gradient analyses on FC data in order to 1) elucidate the principal axes of functional organisation within the LIFG and 2) assess whether there is evidence for graded functional differences.

2. Methods

We used a data-driven approach to extract LIFG gradients based on two measures of FC: 1) correlations in task-free fMRI time-series and 2) meta-analytically derived patterns of task-driven co-activation from across multiple cognitive domains. This bimodal approach not only allowed us to validate our results using independent datasets, it made it possible to assess the generalisability of the functional organisation of the LIFG across different mental states. Indeed, one data type captures activation patterns associated with spontaneous thought (e.g., a state of mind-wandering; Chou et al., 2017; Doucet et al., 2012), while the other is assumed to reflect mental processes

constrained by extrinsic demands (Laird et al., 2013). The summary of our analytical approach is as follows. For each voxel, we 1) extracted BOLD fluctuations over time from resting-state fMRI scans, and 2) meta-analytically identified the brain voxels with which it consistently co-activates across a broad range of task demands. Then, for each FC modality, we compared the fMRI time-series/co-activation patterns of each pair of voxels within the LIFG region of interest (ROI) (see Sections 2.3.1 – 2.3.2). We then conducted gradient analyses on the resulting similarity matrices to extract the principal axes of variation and to estimate the degree of gradation (see Section 2.3.3). In a second step, we conducted descriptive analyses to understand which FC differences gave rise to these gradients (see Section 2.3.4). To this end, we performed seed-based resting-state FC and meta-analytic co-activation modelling (MACM) analyses on hard clusters extracted from the extreme ends of the identified gradients. Finally, we probed the functional/task terms (e.g., "cognitive control", "language") associated with these IFG sub-regions using functional decoding analyses. A schematic overview of the analytic pipeline is illustrated in Figure 1.

2.1. Definition of the LIFG region of interest

The LIFG ROI was created by combining the pars opercularis, pars triangularis, and pars orbitalis as delineated in the second release of the Automated Anatomical Labeling (AAL2) atlas (Rolls et al., 2015). In addition, we included the region termed lateral orbital gyrus in the AAL2 parcellation because it is considered to pertain to pars orbitalis (Keller et al., 2009). These regions correspond roughly to Brodmann areas 44, 45 and (part of) 47. We retained only the voxels with 50% or greater probability of being grey matter according to the ICBM-152 template (Fonov et al., 2009). To ensure the ROI did not encompass regions within neighbouring gyri that were of no interest to the present study, the ROI was manually cleaned by removing voxels that crossed gyral boundaries into the precentral gyrus and middle frontal gyrus in the MNI-152 T1 template included in FSL (version 6.0.1). The final ROI comprised 1,813 (2x2 mm) voxels and is depicted in Figure 1 (step 1) and available at: osf.io/u2834/ .

Analysis Pipeline

Task-based Co-activation   **Task-free Functional Connectivity**

Left Inferior Frontal Gyrus Region of Interest

Step 1:

Extract Data Per Voxel

Identify all studies in NeuroQuery that reported activation in the seed's neighbourhood

ALE Analysis

Whole-brain Co-Activation Map

Extract **BOLD time-series** from pre-processed rs-fMRI data

Spontaneous BOLD Fluctuations

Time

Step 2:

Compute Similarity Matrix

Pairwise comparison of seeds' whole-brain co-activation maps

Similarity Matrix

ROI Voxels

ROI Voxels

Pairwise comparison of seeds' time-series

Participant-level similarity matrix for each run

Z-transform and average across runs & participants

Group Similarity Matrix

ROI Voxels

ROI Voxels

Step 3:

Identify Principal Axes of Variation

Diffusion Embedding

Gradient Maps

Gradient 1   Gradient 2

Diffusion Embedding

Gradient Maps

Gradient 1   Gradient 2

Step 4:

Functional Characterization

Extract voxels located at the extremes of the gradients to use as seeds in further analyses

Hard Clusters

MACM Analyses   Functional Decoding

Task-based Co-activation Maps   **Functional Terms**

Extract voxels located at the extremes of the gradients to use as seeds in further analyses

Hard Clusters

Seed-based Functional Connectivity Analyses

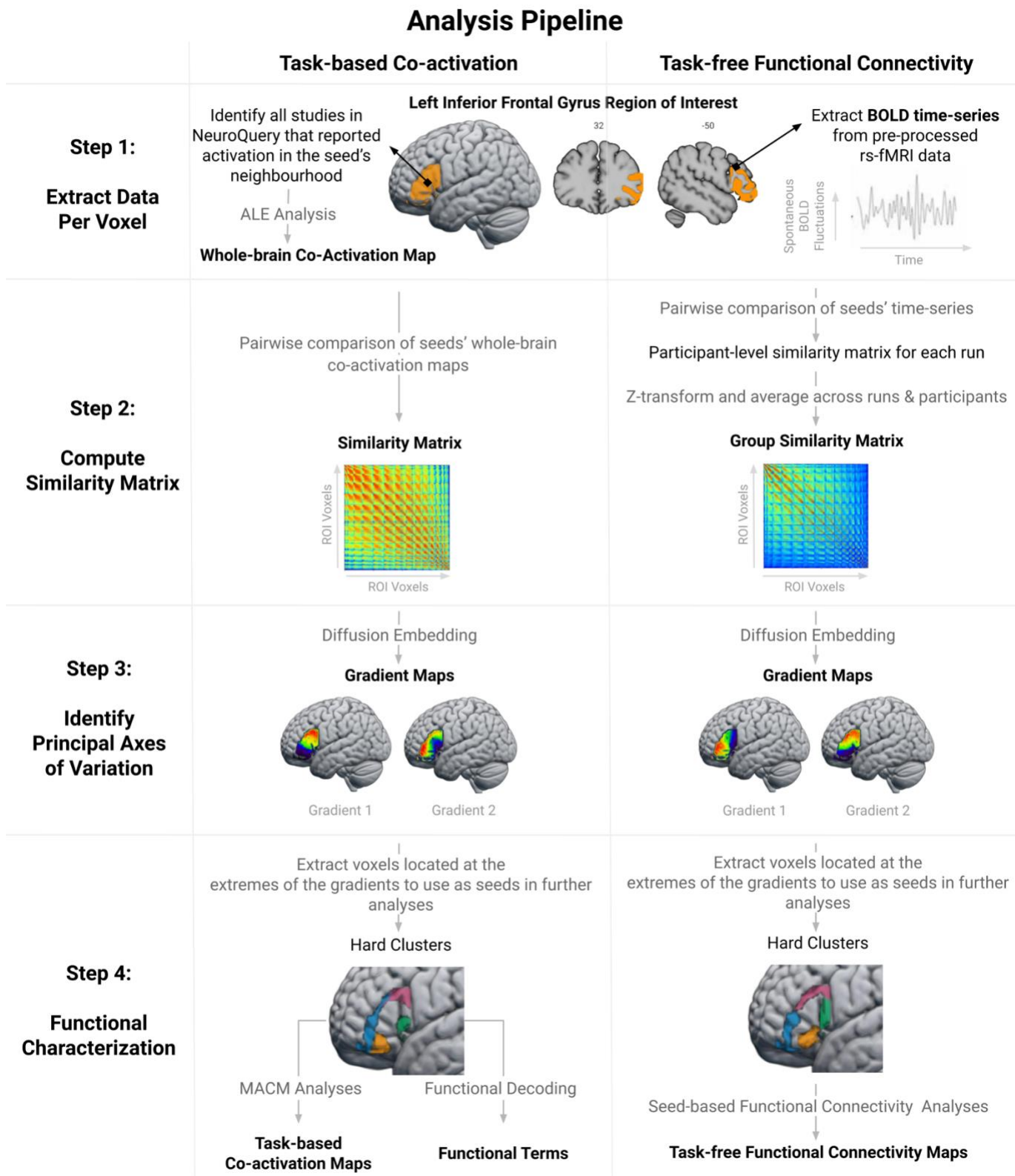**Task-free Functional Connectivity Maps**

Figure 1. Schematic overview of the analytic pipeline. The output of each analysis step is highlighted in bold. In the first step, we estimated the whole-brain co-activation patterns

of individual ROI voxels using meta-analytic co-activation modelling (first column) and extracted their resting-state BOLD time-series (second column). Then, we performed a pairwise comparison of each voxel's co-activation patterns/time-series using the product-moment correlation coefficient. This resulted in two similarity matrices, which were subsequently used as input for gradient analyses (diffusion embedding algorithm) in order to identify the main axes of variation across the ROI. In the final step, we performed MACM, functional decoding and seed-based resting-state FC analyses on hard clusters extracted from the edges of the gradient maps in order to identify the patterns of FC that characterize different IFG sub-regions. The code used for data analysis can be accessed at: osf.io/u2834/.

2.2. Data

2.2.1. Resting-state fMRI data

To assess the functional organisation of the LIFG based on task-free FC, we used the resting-state fMRI time-series of 150 randomly-selected healthy young adult participants (77 females) from the Human Connectome Project S1200 release (Van Essen et al., 2013). For each participant, data were available from up to four 15-minute runs of resting-state fMRI scans collected using the acquisition protocol described by Smith et al. (2013). All four scans were available for 139 participants (92.7% of participant sample), only three scans for three participants (2%) and only two scans for eight participants (5.3%). The data were already pre-processed in MNI space using the minimal processing pipeline described by Glasser et al. (2013) and de-noised using the ICA-FIX approach (Salimi-Khorshidi et al., 2014). We regressed the global signal to further reduce the effects of motion artefacts (Burgess et al., 2016), and smoothed the images using a 4-mm full-width half-maximum Gaussian kernel. In keeping with other resting state studies, we took an additional step of band-pass filtering the data to retain only frequencies between .01 and .08 Hz (Satterthwaite et al., 2013).

2.2.2. Meta-analytic functional neuroimaging data

To assess the functional organisation of the LIFG based on task co-activation patterns, we adopted a meta-analytic approach and capitalized on the openly available NeuroQuery database (neuroquery.org). NeuroQuery contains over 400,000 activation coordinates that were

automatically-extracted from 13,459 neuroimaging studies (Dockès et al., 2020). The database also includes estimates of frequency of occurrence of 6308 terms (e.g., 'cognitive control', 'semantic memory') in each full-text publication from this corpus, which were used to perform functional decoding (see Section 2.3.4).

2.3. Data analysis

2.3.1. Task-free FC similarity matrix

To compute the task-free FC similarity matrix, we first extracted the blood oxygen-level-dependent signal timeseries of every voxel within the ROI, resulting in a voxel by timepoint matrix for each participant and each run. Then, we computed a cross-correlation matrix by calculating the product-moment correlation coefficient between the timeseries of all pairs of ROI voxels. The resulting voxel by voxel matrix was z-score normalized to allow the result of each run to be averaged (Dunlap et al., 2013), in order to generate an average similarity matrix across runs per participant. These participant-level matrices were subsequently averaged resulting in a group similarity matrix. This task-free FC-based similarity matrix was transformed back from z-scores to correlation values for gradient decomposition (see Section 2.3.3).

2.3.2. Task-based co-activation similarity matrix

To compute the task-based co-activation similarity matrix, we first used MACM analyses to identify the brain areas consistently co-activated with each voxel within the ROI. MACM uses meta-analytic data to quantify the co-occurrence of activation between voxels across a broad range of task demands (Laird et al., 2013). This analysis involved extracting all studies in the NeuroQuery database that reported at least one activation peak within 6-mm of a given voxel. Next, we quantified the convergence of activation across the identified experiments using the revised activation likelihood estimation (ALE) algorithm (Eickhoff et al., 2012) as implemented in the Python library NiMARE (Salo et al., 2022). This process was repeated for all voxels within the ROI, resulting in 1,813 unthresholded MACM maps that estimate the strength of co-activation between each ROI voxel and all other brain voxels (ROI voxel by brain voxel matrix). In the second step, we generated a cross-correlation matrix by calculating product-moment correlation coefficient between the MACM map values of each pair of ROI voxels. The resulting task-based co-activation similarity matrix was used as input for the gradient analysis (see Section 2.3.3).

2.3.3. Gradient analysis

We conducted gradient analyses to separately explore the principal axes of variation in task-free FC and task-based co-activation patterns across the ROI. To this end, we first sparsified the similarity matrices by retaining only the top 10% of values row-wise and computed a symmetric affinity matrix using a cosine kernel. The application of this threshold ensures that the results are only based on strong connections, rather than weak and potentially spurious connections (Vos de Wael et al., 2020).

Then, we generated gradient maps by using the diffusion embedding algorithm as implemented in the BrainSpace Python toolbox (Vos de Wael et al., 2020). Diffusion embedding is a type of non-linear dimensionality reduction based on graph theory that describes the high-dimensional connectivity data in terms of distances in a low-dimensional Euclidian space, where the distance between nodes (i.e., voxels) reflects the strength of their connections (i.e. similarity in FC patterns) (for a detailed description, see Coifman & Lafon, 2006). The diffusion embedding algorithm forces voxels with many and/or strong connections closer together and voxels with few and/or weak connections further apart in the embedding space (resulting in gradient maps). We extracted 10 gradients from each modality-specific matrix, but we further interrogate only the first two gradients as they explained considerably more variation in the data compared to the remaining gradients (see Figure S1).

We quantified the degree of gradation in FC changes across the LIFG by estimating the normalised algebraic connectivity of the similarity matrices. This value corresponds to the second largest eigenvalue of the Laplacian of the matrix and represents a descriptive index of how well connected a graph is (Fiedler, 1973). It ranges from zero, which indicates that the graph comprises at least two completely disconnected sub-graphs, to a value of one, which suggests that the graph is characterised solely by graded differences. Thus, the normalised algebraic connectivity of the similarity matrices is indicative of whether the LIFG comprises at least two sharply delineated sub-regions or graded transitions between sub-regions with differences in connectivity/co-activation patterns (Bajada et al., 2020). We note that, while this value is influenced by the smoothing of neuroimaging data, a value much higher than 0 and close to the maximal value possible of 1 is unlikely to be caused only by artificially induced local gradation (Bajada et al., 2017, 2019, 2020; Jackson et al., 2020). We separately estimated the algebraic connectivity of the task-based co-activation matrix and the group task-free FC matrix. In addition, we assessed the

gradation in task-free FC matrices at the participant level. This was done to avoid relying only on a gradation metric derived based on the group matrix which is generated by subjecting the individual-level matrices to an additional transformation that may bias the gradation metric.

2.3.4. Functional characterization

While the gradient analysis can estimate the main directions of functional changes, this step alone cannot reveal the qualitative differences in the FC patterns that drive the functional organisation of the LIFG. To describe the task-free FC and task-constrained co-activation patterns of distinct IFG sub-regions, we defined hard clusters based on their locations at the extremes of the gradients by extracting the voxels with the 20% lowest and highest gradient values. The hard clusters are depicted in Figure 1 (step 4) and their MNI coordinates are reported in Table S1. In a graded map, voxels located at the gradient poles should differ most in terms of their FC patterns. Thus, contrasting the task-free FC and task-constrained co-activation characteristics of the clusters located at the extremes of the gradients allows the identification of the patterns that have driven the separation between the clusters in the embedding space. These clusters should not be interpreted as a hard parcellation of the LIFG.

To identify the FC patterns for each cluster, we used the clusters defined based on each of the task-free FC gradient maps as seeds in seed-based resting-state FC analyses. These analyses were performed using the Python package Nilearn (Abraham et al., 2014). For each participant, we used the average resting-state fMRI time-series (concatenated across runs; Cho et al., 2021) of all voxels within each cluster as a regressor in a general linear model predicting the timeseries of all grey matter voxels. The resulting cluster FC maps were z-transformed and tested for consistency across participants using a one-sample t-test. In addition, to identify the FC specific to each cluster, which is driving the identification of the gradient, paired-samples t-tests were used to generate contrast maps showing the brain regions with greater FC to one hard cluster than the cluster extracted from the opposite end of the same gradient (anterior vs posterior cluster, dorsal vs ventral cluster). The group-level FC maps were thresholded using a family-wise error (FWE) corrected voxel-height threshold of $p < 0.05$ and the probabilistic threshold-free cluster enhancement approach as implemented in the R package pTFCE (Spisák et al., 2019). We wanted to identify the brain regions that i) display greater functional coupling with one LIFG cluster than the cluster at the opposite end of the same gradient and, at the same time, ii) are significantly

coupled with the respective LIFG cluster. Therefore, the contrast maps (determined using the paired-samples t-tests) were masked by the significant connectivity of each cluster (determined using the one-sample t-tests).

To identify the co-activation patterns of each LIFG cluster, we conducted MACM analyses on seeds defined based on each task-based gradient map using the Python package NiMARE (Salo et al., 2022). Specifically, we ran ALE analyses on all studies from the NeuroQuery database that reported at least one activation peak within the seed (see Table S2 for the number of studies identified for each cluster) to identify the brain regions consistently involved in the studies that activate the seed. Specifically, the resulting MACM maps quantify the convergence of activation across all studies that reported activation within the seed. These maps were thresholded using a FWE corrected voxel-level threshold of $p < 0.05$. Then we conducted contrast analyses to identify the brain regions that co-activate more consistently with one hard cluster than the cluster extracted from the opposite end of the same gradient (anterior vs. posterior cluster, dorsal vs. ventral cluster). The contrast maps were thresholded using an uncorrected $p < 0.05$ threshold. To understand which brain regions display (i) greater co-activation with the cluster located at one extreme of the gradient than the other extreme and (ii) significant co-activation with the cluster, we masked the contrast maps by the significant cluster-specific MACM map (determined using independent ALE analysis).

It is important to note that we conducted contrast analyses using the same FC modality (i.e., MACM of clusters extracted from the task-based gradients, seed-based resting-state FC analyses of clusters extracted from the task-free gradients) in order to visualize the differences that have driven the gradients, and not to test whether there were significant FC differences between the clusters. The non-inferential and descriptive nature of these follow-up analyses circumvents analytic circularity (Eickhoff et al., 2015). Nonetheless, we repeated these sets of analyses using an independent FC modality (i.e., MACM of clusters extracted from the task-free gradients, seed-based resting-state FC analyses of clusters extracted from the task-based gradients) to confirm whether the FC maps are consistent regardless of the approach adopted to define the clusters. These analyses revealed similar FC patterns and are only reported in supplementary Figures S8-9.

In line with recent recommendations (Uddin et al., 2022), we determined the network affiliations of our novel findings by comparing them with a commonly-used parcellation scheme. We used the 7-network parcellation proposed by Yeo et al. (2011) as the reference atlas. For each

task-free FC and task-based co-activation map, we computed the percentage of voxels that overlap with each of the seven reference networks. The LIFG has been consistently implicated in semantic processing (Jackson, 2021) which is thought to be supported by a functional network that is dissociable from other canonical networks such as the core default network (DN) (Branzi et al., 2020; Humphreys et al., 2015; Jackson et al., 2016, 2019; Jung & Lambon Ralph, 2022). Therefore, we also computed the overlap between our results and a mask of the semantic network (SN) proposed by Jackson et al. (2016). The reference SN map represents the set of regions that were significantly functionally coupled with the left ventrolateral anterior temporal lobe (ATL), which has been attributed a crucial role in semantic cognition (Binney et al., 2010; Jackson et al., 2016; Lambon Ralph et al., 2017). It is of note that we are not able to dissociate between the DN and SN in these analyses (as has been done by, for example, Humphreys et al., 2015; Jackson et al., 2019) because there is a considerable degree of spatial overlap between the DN mask from Yeo et al. (2011) and the semantic network obtained by Jackson et al. (2016).

Lastly, to identify functional terms associated with each cluster as an index of its potential function, we conducted functional decoding analyses using the BrainMap chi-square approach as implemented in NiMARE (Salo et al., 2022). For each term in the NeuroQuery database, the consistency analysis (also known as forward inference) computes the likelihood of activation reported within the seed given presence of the term in the article's text, whereas the specificity analysis (also known as reverse inference) estimates the posterior probability of an article containing the term given activation reported inside the seed. The results of these analyses were thresholded at $p < 0.05$ using the Benjamini-Hochberg false discovery rate correction. To aid the interpretability of the results, we retained only the terms with at least 80% likelihood of being related to cognitive functions based on rater annotations (Bottenhorn et al., 2018).

3. Results

3.1. Gradient maps

The first two task-free FC gradients were selected for further analysis because together they accounted for > 50% of variance, while the lower-order gradients explained less than 11% of variance each (Figure S1). The voxels' gradient values, which reflect the similarity between their resting-state fMRI timeseries, were visually coded and projected on the brain using a colour spectrum from red to dark blue to reveal the pattern of change in task-free FC across the LIFG. As

can be seen in Figure 2A, the FC patterns of the LIFG are principally organized along an anterior-posterior axis that accounted for 30% of the variance. This gradient progressed from the anterior portion of the LIFG, bordering the inferior part of the inferior frontal sulcus (IFS), to the posterior region, bordering the precentral gyrus. The second gradient, which explained 25% of the variance, revealed changes in connectivity along the superior-inferior dimension. This gradient progressed from the superior part of the IFS and the precentral sulcus to the inferior portion of the IFG, bordering the lateral orbital sulcus. The algebraic connectivity of the group similarity matrix was 0.71, suggesting a high level of gradation in task-free FC changes across the LIFG. This was confirmed by the distribution of the algebraic connectivity of the individual-level similarity matrices (Figure S2) which had a mean of 0.89 (SD = 0.02). The group similarity matrix, reordered based on the voxels' positions along the first and second gradients, and showing the graded change in FC across voxels in the LIFG, is illustrated in Figure S3.

The first two task-based co-activation gradients were selected for further analysis because together they accounted for > 60% of the variance, while the lower-order gradients individually explained less than 11% of the variance (Figure S1). The principal gradient accounted for 42% of the variance and progressed along a dorsal-ventral axis from the inferior frontal junction (IFJ) to the antero-ventral region bordering the lateral orbital sulcus and inferior portion of the IFS. The second gradient explained 21% of the variance and revealed changes in connectivity that followed the rostral-caudal axis in a radial pattern progressing from the inferior portion of the pars opercularis towards the IFS. The algebraic connectivity of the co-activation similarity matrix was 0.77, suggesting that LIFG is characterized by gradual changes in consistent patterns of co-activation across cognitive domains (see Figure S3 for the reordered matrices). Because the unit of the task-based analysis is the study rather than the participant, the gradation cannot be assessed at the participant level as in the case of the task-free analysis reported above.

The gradients extracted from the two independent datasets converge on two principal organisational axes of the LIFG: anterior-posterior and dorsal-ventral. Visual inspection of the gradient maps suggests that the first task-free gradient and the second task-based gradient capture a similar anterior-posterior axis of functional variation, which is supported by a strong positive correlation of 0.77 between voxels' position ranks on the two gradients (see Figure S4 for the scatterplot). Likewise, the second task-free gradient and the first task-state gradient capture a similar dorsal-ventral organisational dimension. This observation is supported by a strong positive

correlation of 0.7 between voxels' position ranks on the two gradients (Figure S4). The orders in which these gradients appear are switched between the task-free and task-constrained FC data, and this is because of a difference in the relative amount of variance explained by each gradient. Because it is subtle relative to the similarities, this difference could be attributable to noise but it may also reflect meaningful differences in the connectivity revealed by task-free and task-constrained mental states (Eickhoff & Grefkes, 2011).
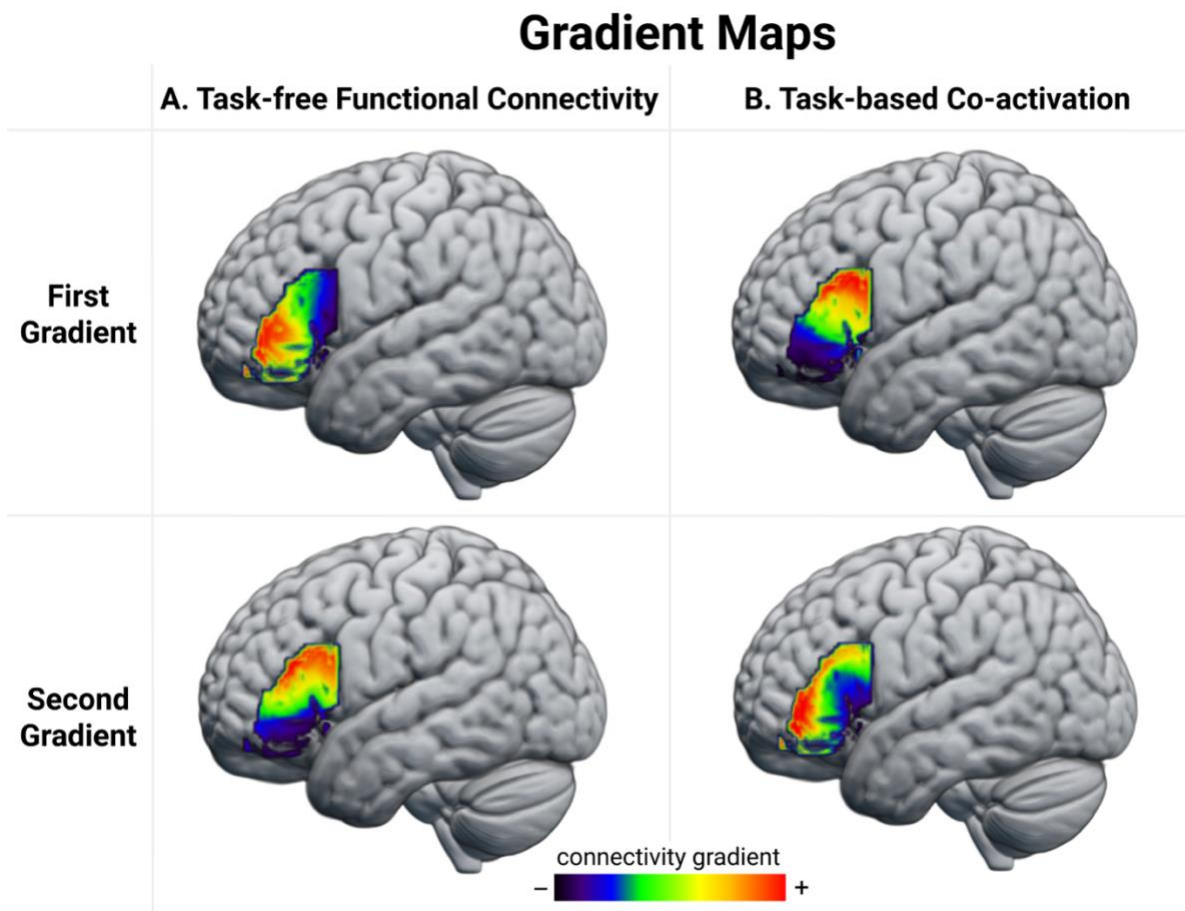
Figure 2. The first two gradient maps extracted from the A) task-free FC similarity matrix and B) task-based co-activation similarity matrix. Compared to regions represented with colours further apart on the colour spectrum, regions represented using colours that are closer together show greater similarity in their A) correlation with each other over time during resting fMRI scans and B) their patterns of co-activation across tasks spanning a range of cognitive domains. The +/-indicate different poles of these gradient dimensions, but the assignment to a specific end of a dimension is arbitrary. The gradient maps can be accessed at: neurovault.org/collections/ETPEFCDV/.

3.2. Functional characterization

3.2.1. Differential task-free FC patterns

We contrasted the whole-brain resting-state connectivity patterns of the clusters located at the extremes of the anterior-posterior task-free gradient. This revealed differences in their functional coupling with a bilateral and distributed set of brain regions (Figure 3A; Table S3). The anterior cluster showed stronger FC with frontal regions, including the right IFG (pars orbitalis), bilateral IFS, dorsal and orbital portion of the middle frontal gyrus (MFG), superior frontal gyrus (SFG), medial prefrontal cortex (mPFC), and orbitofrontal cortex (OFC), with parietal regions in the posterior cingulate cortex (PCC), angular gyrus (AG), and inferior parietal lobule (IPL), and with temporal regions along the length of the middle and inferior temporal gyri (MTG/ITG) and in the fusiform gyrus (FG), and left hippocampus. In contrast, the posterior cluster showed stronger FC with frontal regions in the right IFG (pars opercularis), the middle portion of the MFG, precentral gyrus, pre-supplementary motor area (pre-SMA), anterior and middle cingulate cortex (ACC), and insula, and with posterior cortical regions in the supramarginal gyrus (SMG), and posterior superior temporal sulcus, and with basal ganglia.

Comparison between the task-free FC of the dorsal and ventral clusters revealed stronger coupling between the dorsal LIFG and frontal regions in bilateral IFS and IFJ, left MFG, and pre-SMA, parietal cortex in bilateral IPS, and left IPL, and temporal cortex in bilateral posterior ITG and left FG (Figure 3A; Table S4). In contrast, the ventral LIFG showed increased connectivity to the frontal cortex in the right IFG (pars triangularis and pars orbitalis), bilateral SFG, mPFC, and ACC, to the precuneus, a swathe of temporal cortex progressing from the bilateral ventrolateral ATL through the MTG towards the AG, and to the left hippocampus.

Comparison between the cluster-specific task-free FC patterns and canonical networks indicate stronger functional coupling between the anterior LIFG and regions falling within the bounds of the SN/DN and the frontoparietal network (FPN), and between the posterior LIFG and brain regions associated with the ventral attention network (VAN) and somatomotor network (SMN). The dorsal LIFG showed stronger FC with regions of the FPN and dorsal attention network (DAN), whereas the ventral LIFG showed a preference for SN/DN regions. Additional conjunction analyses showed that both the anterior and posterior clusters are coupled with regions of the FPN and SN/DN, and that both the dorsal and ventral clusters are functionally connected mainly with SN/DN regions (Figure S7A; Table S3-4).

3.2.2. Differential task-constrained co-activation patterns

The anterior LIFG showed increased consistent co-activation across a wide variety of tasks with frontal regions in the right IFS and precentral gyrus, and also with bilateral IPS and left posterior ITG, whereas the posterior LIFG co-activated more with the right IFG, bilateral anterior insula and left superior temporal gyrus (Figure 3B; Table S5). The dorsal cluster co-activated more with frontal cortex in the right IFJ, bilateral precentral gyrus, dorsal anterior insula, and pre-SMA, and with the IPS, and left posterior FG (Figure 3B; Table S6). In comparison, the ventral cluster showed increased co-activation with the right IFG (pars orbitalis), and left mPFC, MTG and AG. Given the conservative threshold applied to the independent maps, we also looked at the whole contrast maps without masking by these independent maps. These additionally revealed more consistent co-activation of the posterior cluster with the bilateral STG and of the ventral cluster with the bilateral ATL, precuneus and left AG (Figure S6).

Comparison between the cluster-specific task-based co-activation patterns and canonical networks shows that the anterior LIFG cluster co-activates more consistently with brain regions that are part of the FPN and DAN, whereas the posterior LIFG cluster co-activates mainly with regions associated with the VAN and, when additional masking is not applied, the SMN. The dorsal cluster co-activates preferentially with regions of the FPN and DAN, whereas the ventral LIFG cluster shows stronger co-activation with the DN/SN. Additional conjunction analyses showed overlap between the co-activation maps of the anterior and posterior clusters and those of dorsal and ventral clusters primarily in regions of the FPN (Figure S7B; Table S5-6).

The FC analyses performed on clusters extracted from the gradient maps derived using the independent dataset (i.e., seed-based FC analyses of clusters derived using NeuroQuery studies and MACM analyses of clusters derived using task-free fMRI timeseries), which were conducted to assess the robustness of the results across different strategies for defining seeds, revealed a similar pattern of results (Figure S8-9).

3.2.3. Comparison between the task-free and task-based FC patterns

The task-free and task-based analyses implicate overlapping regions, although the clusters identified in the task-based analyses were less extensive. Specifically, the anterior LIFG was connected with executive control regions (e.g., IFJ, IPS; Assem et al., 2020; Camilleri et al., 2018; Fedorenko et al., 2013), but in the task-free maps it was also connected to regions implicated in

semantic cognition (e.g., ATL, AG; Binder & Desai, 2011; Lambon Ralph et al., 2017). Further, the posterior LIFG was connected to areas that have been ascribed important roles in sensorimotor processing, as well as in phonological and articulatory linguistic processes (e.g., bilateral STS/STG, but in the task-free maps it was also connected to motor and premotor cortices, SMA, MFG and SMG; Hartwigsen et al., 2010; Hickok, 2009; Hickok & Poeppel, 2007; Price, 2012; Ueno et al., 2011; Vigneau et al., 2006). This cluster was also connected to regions considered crucial for salience processing (e.g., anterior insula, but in the task-free results also to dorsal ACC; Menon & Uddin, 2010; Uddin, 2015). The dorsal LIFG was connected to regions that are implicated in executive function (e.g., IFJ, MFG, IPS; Assem et al., 2020; Camilleri et al., 2018; Fedorenko et al., 2013), whereas the ventral LIFG was connected with a set of regions ascribed key roles in semantic and episodic memory (e.g., ATL, medial temporal lobe, AG; Binder & Desai, 2011; Lambon Ralph et al., 2017)

Despite the similarities in the regions implicated, there were some differences in the network affiliations derived from the task-free and task-based analyses. However, comparing the network affiliations of the different contrast maps directly is not possible because 1) the overlap index depends on the size of the maps, which differs considerably between the task-free and task-based analyses and 2) there are differences between the seeds upon which the task-free and task-based analyses are based (see Figure 1; e.g., the task-based anterior seed extends across the length of the IFS and overlaps with the dorsal LIFG seed, whereas the anterior seed used for the task-free analysis does not). Therefore, we will focus the interpretation on the similarities.

The dorsal LIFG connected to FPN and DAN regions, two networks that contribute to the task-general multiple demand network (MDN; Assem et al., 2020; Majerus et al., 2018). In contrast, the ventral LIFG was affiliated mainly with the DN/SN. The DN and SN cannot be distinguished in our assessment given the high degree of spatial overlap between the masks used. However, we note that both the task-free and the unmasked task-based results suggest strong coupling with the ATL, a key hub of semantic knowledge (Lambon Ralph et al., 2017), as well as with the left hippocampus/parahippocapal gyrus, known to be important for episodic memory (Burgess et al., 2002; Dickerson & Eichenbaum, 2010). As such, the dorsal-ventral organisational dimension seems to distinguish between domain-general control networks at the dorsal end and memory-related networks at the ventral end.

The posterior LIFG showed a preference for the VAN, suggestive of a role in perceptually-driven cognition (Corbetta et al., 2008; Corbetta & Shulman, 2002). The anterior LIFG showed a preference for regions that overlap with the FPN, consistent with a role in cognitive control (Assem et al., 2020). The task-free data revealed additional strong coupling with regions that are part of the DN/SN. The task-based analyses might have led to less extensive association with the DN because this network is known for its tendency to deactivate in response to various task demands (Buckner et al., 2005; Mazoyer et al., 2001; Shulman et al., 1997), but it tends to activate during mind wandering states which frquently occur during resting-state scans (Smallwood et al., 2021). Nonetheless, there is evidence that the DN works with the FPN in support of some types of goal-directed cognition (Spreng et al., 2010, 2014) and that it contributes to cognitive control (Crittenden et al., 2015). As such, the anterior-posterior organisational dimension seems to distinguish between higher-order transmodal networks at the anterior edge and perceptually-driven networks at the posterior edge.

A. Differential Task-Free Functional Connectivity Patterns

Contrast Maps | % Overlap with Canonical Networks

Anterior-Posterior Gradient — anterior > posterior; posterior > anterior

Dorsal-Ventral Gradient — dorsal > ventral; ventral > dorsal

B. Differential Task-Based Co-activation Patterns

Contrast Maps | % Overlap with Canonical Networks

Anterior-Posterior Gradient — anterior > posterior; posterior > anterior

Dorsal-Ventral Gradient — dorsal > ventral; ventral > dorsal

Figure 3. Results of contrast analyses between A) the task-free connectivity patterns (derived using seed-based resting-state FC analyses) of the IFG clusters located at the extremes of the anterior-posterior and dorsal-ventral task-free gradients and B) task-constrained co-activation patterns (derived using MACM analyses) of the IFG clusters located at the extremes of the anterior-posterior and dorsal-ventral task-based gradients. These

contrast maps were masked using cluster-independent maps. The spider plots in the right column show the percentage of voxels in each contrast map that overlap between with each of the seven canonical networks from Yeo et al. (2011), as well as the SN from (Jackson et al., 2016) (available at: github.com/JacksonBecky/templates). Note that percentage values are relative to the size of each contrast map; therefore, only the relative patterns of overlaps within each contrast map are of interest and direct comparisons between the network affiliations of different contrast maps are misleading. The contrast maps can be accessed at: neurovault.org/collections/ETPEFCDV/. Abbreviations: SMN - somatomotor network, VAN – bentral attention network; DAN – dorsal attention network; FPN – frontopariental control network; DN – default network; SN – semantic network.

3.2.4. Functional decoding

The functional decoding results suggest possible functional associations of the different LIFG clusters. All four LIFG clusters were significantly associated with terms related to semantic and linguistic processing, including *language, semantic, retrieval, reading, and phonological.* Compared to the posterior cluster, the anterior cluster was associated with the terms *executive (function)* and *memory retrieval.* In contrast, the posterior cluster was associated with terms related to perceptual and motor processing, such as *movement, recognition, auditory,* as well as speech-related terms, such as *phonetic* and *vocal.* Compared to the ventral cluster, the dorsal cluster was associated with terms related to a wide range of cognitive/behavioural domains and input modalities, including *visual, auditory, visuospatial, working memory, executive, social, reward, mood.* In contrast, the ventral cluster was associated with terms such as *memories, mentalizing, reappraisal and autobiographical,* which are suggestive of the purported internally-oriented functions of the DN (Smallwood et al., 2021). While functional decoding approaches can provide pointers to the potential task associations of these regions, it is important to note that the specificity of the results is limited due to the limitations of automated data mining tools like NeuroQuery, such as the aggregation of all contrasts reported in an article, regardless of the cognitive aspects they isolate (Dockès et al., 2020). As a consequence, interpretation should focus on the overall patterns that emerge, rather than the associations of individual terms. Detailed lists of the

functional associations are presented separately for forward and reverse inference analyses and task-free and task-based clusters in Supplementary Figures S10-11.

Figure 4 summarizes the functional decoding results that were consistent for the clusters extracted from the task-free and task-based gradients (e.g., terms associated with both the anterior edge of the task-free gradient and the anterior extreme of the task-based gradient). It also includes a schematic of the proposed functional organisation, which takes into account the results of the FC contrast analyses, the network affiliations, the functional decoding, as well as previous literature reviewed in detail in the Discussion.

4. Discussion

The present study made the first attempt to use data-driven gradient analyses of FC data to elucidate the functional organization of the LIFG. We specifically aimed to 1) map the principal axes of change in function, and 2) determine whether these shifts might be graded. In the following two sections, we shall summarise our novel findings, and then discuss their functional significance.

4.1. Graded topographical organisation of the LIFG along two principal axes

Our analyses converged upon two key findings. First, the FC across the LIFG is principally organized along two orthogonal axes. One of these axes is oriented in an anterior to posterior direction and driven by stronger coupling with the FPN and DN in the rostral aspect, and with the VAN and SMN at the caudal end. The second arose along a ventral to dorsal orientation, and reflected greater connectivity of ventral LIFG to the DN, whereas dorsal regions abutting the IFS/IFJ were more tightly coupled with the FPN and DAN. These differential patterns of FC are in line with previous investigations (Barredo et al., 2016; Clos et al., 2013; Davey et al., 2016; Jakobsen et al., 2016, 2018; Kelly et al., 2010; Muhle-Karbe et al., 2016; Neubert et al., 2014; Wang et al., 2020) and suggest that the LIFG interfaces between distinct large-scale functional networks, consistent with its proposed role as a cortical hub (Buckner et al., 2009; Sepulcre et al., 2012).
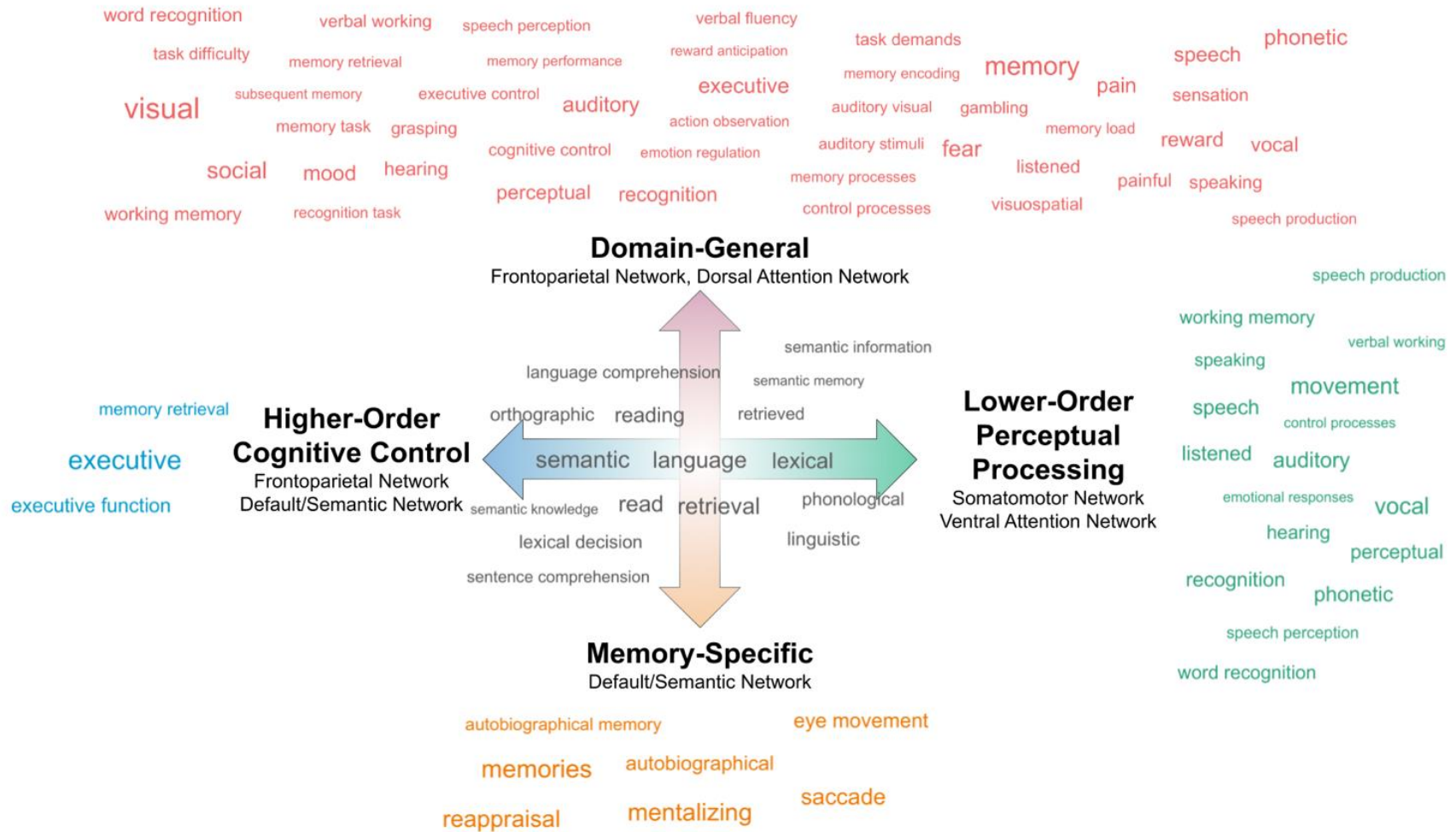
Figure 4. Schematic representation of the proposed functional organisation of the LIFG. The anterior-posterior organisational axis (represented by the horizontal blue-to-green arrow) might reflect a shift from lower-order perceptual processing (posterior LIFG) via affiliation with the SMN and VAN, and higher-order cognitive control (anterior LIFG) via affiliation with the FPN and DN.

The dorsal-ventral axis (represented by the vertical orange-to-red arrow) might reflect a shift from domain-general executive functions (dorsal LIFG) via affiliation with the FPN and DAN and the cognitive control of information stored in long-term memory (ventral LIFG) via affiliation with the DN/SN. The word clouds illustrate functional terms associated with the LIFG clusters located at the extremes of the task-free and task-based gradients in the forward or reverse inference functional decoding analyses - terms associated with the anterior but not posterior cluster (blue); the posterior but not anterior cluster (green); the dorsal but not ventral cluster (red); the ventral but not dorsal cluster (orange); and with all four clusters (grey). The size of the word reflects the effect size of the association.

Our second key finding is that FC of the LIFG shifts in a relatively graded manner. Precisely, the algebraic connectivity of the similarity matrices revealed that FC analyses do not support there being abrupt boundaries and discrete functional parcels in LIFG. This is consistent with both contemporary descriptions of LIFG connectivity based on intraoperative cortico-cortical evoked potentials (Nakae et al., 2020) and structural properties (Binney et al., 2012; Thiebaut de Schotten et al., 2017), as well as classical descriptions that include a fan-shaped set of anatomical projections emanating from the IFG into the lateral temporal lobe (Dejerine & Dejerine-Klumpke, 1895). The present study is the first to confirm a graded organisation of the LIFG using a bimodal dataset, taking into account task-driven variation on one hand, and task-free FC on the other.

Overall, our findings are compatible with previous parcellations despite key differences in the methodological approaches (Clos et al., 2013; Kelly et al., 2010; Klein et al., 2007; Neubert et al., 2014; Wang et al., 2020). This includes those that have taken a 'hard' clustering approach to parcellating left prefrontal cortex (PFC). For example, Neubert et al. (2014) parcellated PFC based on structural connectivity and found that the LIFG fractionated into discrete subdivisions positioned along the anterior-posterior dimension. The connectivity of these parcels was distinct from those situated dorsally in the adjacent IFS and IFJ, which implies a further dorsal-ventral dimension of organisation. The co-existence of these two axes of LIFG organisation is also apparent in hard parcellations of the LIFG (Clos et al., 2013; Wang et al., 2020), its right hemisphere homologue (Hartwigsen et al., 2019), and more encompassing parcellations of cortex (Glasser et al., 2016). Of course, the results of graded and hard parcellation are not identical as

hard parcellations: (i) force voxels that are part of intermediate regions with gradually-changing connectivity to be within the borders of discrete clusters (Bajada et al., 2017; Haueis, 2012) and (ii) require the a priori specification of the number of clusters that are to be identified, perhaps making them insensitive to finer details. However, the two superimposed yet orthogonal modes of organization identified here may have driven prior hard parcellations of the LIFG.

4.2.   The putative functional significance of the LIFG's FC gradients

Taken together, the cluster-specific FC patterns and functional decoding results paint a coherent picture regarding the functional significance of the graded connectivity patterns that appear across the LIFG. On this basis, and in conjunction with the results of previous functional neuroimaging studies, we propose the following interpretation, which has also been illustrated schematically in Figure 4. First, the dorsal-ventral axis might reflect a functional transition from domain-general executive function (dorsal LIFG) to domain-specific control of meaning-related representations (ventral LIFG). Second, the anterior-posterior axis might reflect a shift from perceptually-driven processing (posterior LIFG) to higher-level transmodal control (anterior LIFG). We discuss this proposal in further detail below.

The dorsal LIFG was functionally coupled with regions that comprise the FPN and DAN. These two networks contribute to a wide variety of task demands that span multiple cognitive domains (Assem et al., 2020; Cole et al., 2013). In contrast, the ventral LIFG was preferentially affiliated with the DN, as well as brain regions that have been ascribed key roles in semantic cognition, such as the anterior temporal lobes (Binney et al., 2010; Lambon Ralph et al., 2017). Thus, the shift in FC towards ventral IFG subregions might reflect a specialization towards the application of cognitive control to prior knowledge. Indeed, it has been proposed that the LIFG, as a whole, sits in a unique position at the intersection of the MDN and the DN, and that this makes it ideally suited for implementing demanding operations on meaning-related representations (Chiou et al., 2022; Davey et al., 2016). Consistent with this, the LIFG responds reliably to an increased need for the control of semantic information across a wide range of experimental paradigms (Diveica et al., 2021; Jackson, 2021), including those requiring episodic memory retrieval (Vatansever et al., 2021). However, it is increasingly apparent that there are finer-grained functional distinctions within the LIFG; dorsal LIFG regions near IFS/IFJ overlap with the MDN and are engaged by control demands that are common across many cognitive tasks/domains

(Assem et al., 2020; Fedorenko et al., 2013; Hodgson et al., 2021), which may include phonology (Hodgson et al., 2021; Poldrack et al., 1999), whereas the ventral LIFG contributes selectively to challenging semantic tasks (Gao et al., 2021; Whitney et al., 2011, 2012). One possible explanation is that ventral LIFG is specifically involved in controlled semantic retrieval processes as opposed to domain-general selection mechanisms, which are under the purview of dorsal LIFG regions (Badre & Wagner, 2007; Barredo et al., 2015; but see Snyder et al., 2011; Whitney et al., 2012). Alternatively, the processes implemented might be equivalent, but connectivity differences mean that they operate on distinct sets of inputs/outputs.

The anterior and posterior LIFG clusters were each affiliated with networks that occupy different positions along a macroscale cortical hierarchy that transitions from sensorimotor to transmodal cortex (Margulies et al., 2016). Specifically, the posterior LIFG was connected with the VAN and SMN, which process inputs from the external environment (Corbetta et al., 2008; Menon & Uddin, 2010). In contrast, anterior LIFG was preferentially coupled with regions of the FPN and DN, which are positioned towards the top end of the cortical hierarchy (Margulies et al., 2016). The anatomical and functional separation of anterior LIFG regions from sensorimotor systems might be requisite for the implementation of perceptually-decoupled, temporally-extended, and higher-order cognitive control (Fuster, 2001; Kiebel et al., 2008; Raut et al., 2020; Taylor et al., 2015). This interpretation is consistent with the proposal that the PFC is characterized by a posterior-anterior gradient of hierarchical control (for a review, see Badre & Desrochers, 2019), which was motivated by studies showing that anterior PFC is preferentially engaged by tasks that require generalization over an extended set of rules, integration of a larger number of dimensions and/or contexts sustained over longer periods of time (Badre & D'Esposito, 2007; Bahlmann et al., 2015; Koechlin et al., 2003; Nee & D'Esposito, 2016; for an investigation focused on the LIFG, see Koechlin & Jubault, 2006). In the language domain, it has been suggested that LIFG has a key role in the integration of linguistic subordinate elements into superordinate representational structures, and that this reflects a caudal-rostral functional gradient from phonological to syntactic to conceptual processing (Hagoort, 2005; Uddén & Bahlmann, 2012; also see Asano et al., 2021; Jeon & Friederici, 2015).

4.3.   Concluding remarks

Our analyses revealed two main axes of organisation in LIFG function, in anterior-posterior and dorsal-ventral orientations, which is consistent with broader proposals concerning the whole PFC (Petrides, 2005). Moreover, our results suggest that functional differentiation across the LIFG occurs in a graded manner, and we were not able to find any clear evidence for discrete functional modules. Crucially, we replicated the principal gradients using two independent measures of FC, which suggests that our results are not dependent on idiosyncrasies of the datasets, and instead reflect stable, generalizable properties of LIFG organisation. The high degree of cross-modal similarity also suggests that a comparable LIFG functional organisation underpins divergent mental states. Future work is needed to directly probe the functional significance of these organisational dimensions and assess the compatibility of our findings at different spatial scales (e.g., cellular) and within other neuroimaging modalities (e.g., tractography) such that it is possible to arrive at an integrated account of the functional organisation of the LIFG.

5.   Extended results/discussion: implications for our understanding of social cognition

The previous sections of this chapter forms part of a manuscript intended for submission for peer-review and publication. Below I shall provide some supplementary results and discussion, which are aligned with the broader objectives of this thesis.

5.1.   Projection of social meta-analytic activation maps onto the LIFG gradients

The main aim of this additional analysis was to establish where social cognition is situated on the principal LIFG gradients. The LIFG gradients capture the two spatial dimensions that explain most variance in FC patterns with the rest of the brain and are indicative of functional variation within the LIFG. Therefore, the LIFG gradients can be used as topographical maps that (i) might provide clues about the role of the LIFG in social information processing and (ii) might help better understand the similarities and differences between the neural correlates of different social cognitive tasks.

The meta-analytic maps of the five social sub-domains studied in Chapter 4, specifically theory of mind (ToM), trait inference, empathy for pain and emotions, and moral reasoning, were projected onto the principal LIFG gradients. To accomplish this, the LIFG gradient maps derived from the task-free data were divided into equally-sized decile bins (the voxels with the top 10%

highest values were assigned to bin 1, next highest 10% of values to bin 2 and so on). For each meta-analytic map, the average activation likelihood value of all voxels in each bin was computed and plotted in Figure 5. This analysis was also performed on meta-analytic maps of semantic control manipulations from Diveica et al. (2021), and n-back working memory manipulations from Hodgson et al. (2021). This allowed comparing the activation of the LIFG in response to social tasks with activation sensitive to two different cognitive control manipulations.
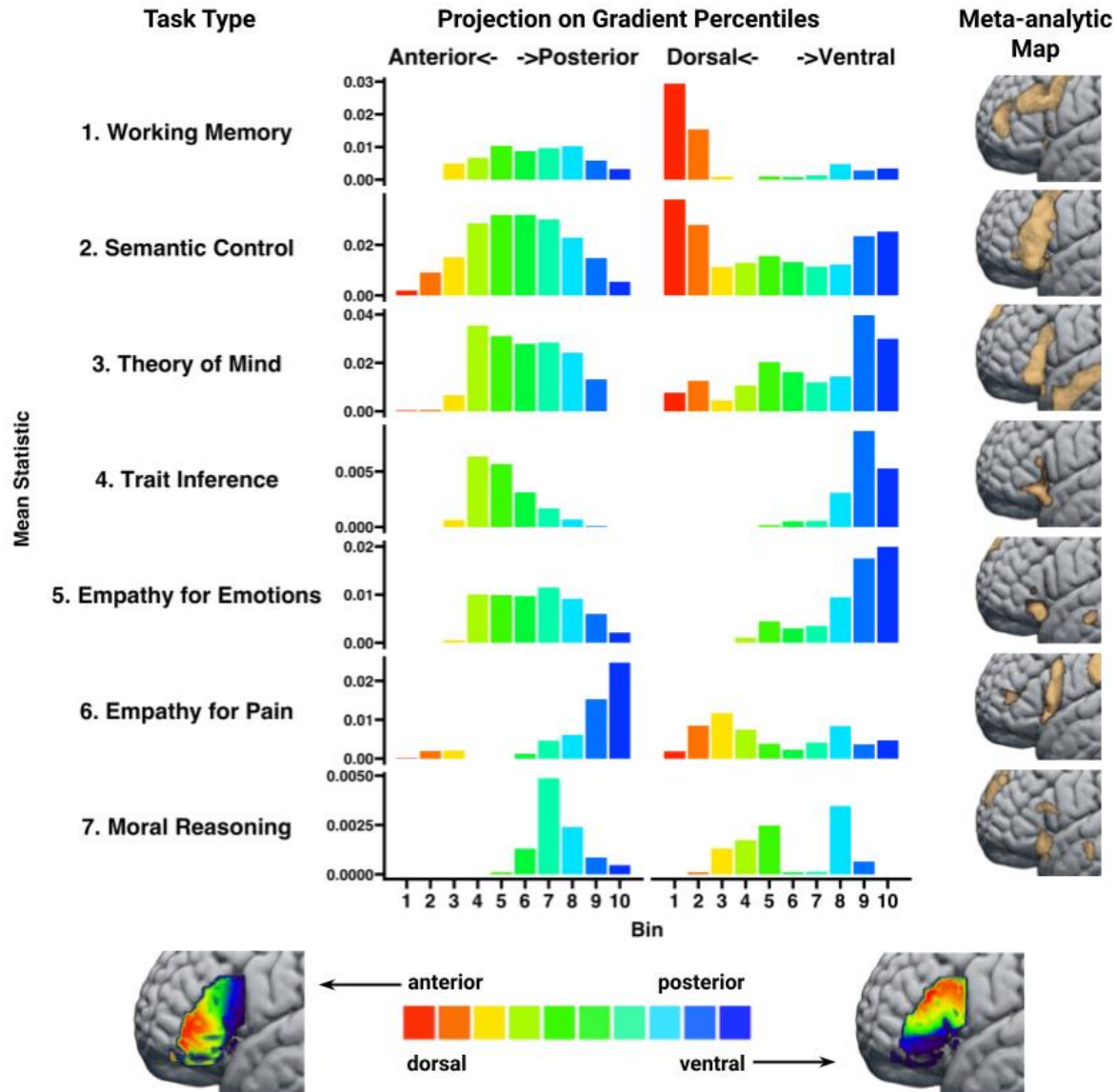

Figure 5. The likelihood of activation in response to social tasks and executive control manipulations projected onto the principal LIFG gradients derived from task-free FC data. The mean activation likelihood statistic is plotted on the y axis; the higher the mean

statistic, the more likely that the gradient bin activates in response to the cognitive domain specified. The bin number is plotted on the x-axis and progresses from anterior LIFG to posterior LIFG in the first column, and from the dorsal LIFG to the ventral LIFG in the second column. The colours emphasize the rough correspondence between the bin number and the associated range of gradient values. The right-most column illustrates the FWE-corrected meta-analytic maps of the social/cognitive abilities, which quantify the consistency of activation across functional neuroimaging studies.

5.2. Dissociable contributions of the dorsal-ventral LIFG gradient to social cognition

Our findings are in line with the proposal that the dorsal and ventral LIFG make dissociable contributions to social (and non-social) cognition. As can be seen in Figure 5, working memory and semantic control manipulations are associated with differences in the distribution of activation across the dorsal-ventral LIFG gradient. Specifically, the dorsal LIFG is reliably recruited by both control manipulations, indicative of a task-general control function and consistent with its strong coupling with domain-general control networks. Indeed, there is direct evidence to suggest that the dorsal LIFG contributes to social cognition via a role in task-general control. For example, the dorsal LIFG was found to be recruited by both ToM and response selection tasks (Saxe et al., 2006) and has also been implicated in working memory for various types of social information, like names (Meyer et al., 2015), traits (Meyer et al., 2012, 2015), faces (Druzgal & D'Esposito, 2003; Roth et al., 2006) and voice identities (Rämä et al., 2001; Rämä & Courtney, 2005; Relander & Rämä, 2009).

In contrast, ventral aspects of the LIFG are preferentially activated by semantic control manipulations, suggestive of a relative specialization for demanding semantic operations (also see Chiou et al., 2022; Gao et al., 2021). This interpretation is consistent with its strong coupling with DN/SN regions, including the ATL semantic hub. Although it is unclear whether the ventral LIFG responds to semantic control manipulations within complex social tasks, this region has been shown to support the retrieval of weakly-associated social concepts (Satpute et al., 2014), in line with the proposal that it enables the controlled retrieval of conceptual knowledge in the service of social cognition. Furthermore, it is this ventral portion of the LIFG that appears to be recruited, to some extent, by all four social abilities, consistent with the view that controlled semantic retrieval

processes are key to social cognition (Binney & Ramsey, 2020; Diveica et al., 2021; Souter et al., 2021).

5.3. Dissociable contributions of the anterior-posterior LIFG gradient to social cognition

Most social abilities investigated and both types of cognitive control demands are associated with convergent activation that spans the full extent of the anterior-posterior LIFG gradient, peaking in its midpoint (Figure 5). The fact that these regions are associated with variation in FC suggests that they might contribute differently to task performance. Indeed, functional dissociations along the anterior-posterior LIFG axis have been reported during social tasks, particularly in the domain of action understanding (Kilner, 2011). Specifically, the posterior LIFG seems to respond to concrete features of actions, such as how an action is performed, whereas the anterior LIFG tends to respond to more abstract action features, like the intentions behind actions (Spunt & Lieberman, 2012). Relatedly, the area at the border of the posterior LIFG and the premotor cortex are part of the mirror neuron network, contributing to both the execution and perception of actions, whereas more anterior LIFG regions do not (Kilner et al., 2009; Press et al., 2012). Thus, it has been suggested that there is a posterior-anterior LIFG gradient that shifts from 'pragmatic' action representations to abstract representations of actions, including goals and intentions (Kilner, 2011). This view is in line with our FC findings, and with the abstraction gradient proposed to characterize the PFC (Badre, 2008; Badre & Desrochers, 2019).

In contrast to the other social abilities, empathy for pain elicits most convergent activation in the posterior portion of the LIFG, which is functionally coupled with the VAN (also known as the salience network). The recognized engagement of this network in empathy for pain is thought to reflect the saliency of the stimuli (i.e., others in painful situations) and a higher need for bottom-up attentional processes that detect and orient attention towards salient events (Betti & Aglioti, 2016; Iannetti et al., 2013). This is in line with the interpretation that posterior LIFG supports perceptually-driven cognitive processes.

Overall, our findings suggest that social abilities recruit areas of the LIFG characterized by different FC with large-scale functional networks. This is in line with the view that different forms of social cognition draw flexibly on the cognitive resources underpinned by distinct neural regions/networks to meet situational demands (Alcalá-López et al., 2018; Binney & Ramsey, 2020;

Schaafsma et al., 2015; Schurz et al., 2020). However, future research is needed to more clearly delineate the nature of the dissociable contributions made by the LIFG to social cognition.

CHAPTER 6

General Discussion

To date, social neuroscience has focused primarily on identifying neurocognitive systems that are tuned for processing social stimuli and tasks, and has paid less attention to the role of domain-general systems (Amodio, 2019; Binney & Ramsey, 2020; Ramsey & Ward, 2020). This thesis seeks to further understanding of the contributions of domain-general systems and takes a novel primary systems approach that positions social information processing within a broader semantic cognition framework (Binney & Ramsey, 2020). This generalist model was chosen because of three key strengths: 1) by framing social cognition as one way in which meaning is gathered from our environment, it can potentially capture a more complete picture of the various neural systems implicated - from sensation to controlled cognition and behaviour; 2) it could offer a unifying model of social cognition that spans multiple behavioural phenomena; and 3) it puts forth several falsifiable hypotheses that are based on decades of multi-method basic science findings as well as clinical observations (Binney & Ramsey, 2020; Lambon Ralph et al., 2017). This chapter summarizes the main findings of this thesis, discusses their implications for neurocognitive models of social cognition, and concludes by suggesting promising avenues for future research.

1. Summary of findings

Chapter 3 sought to address two key methodological limitations of extant research on the role of socialness in conceptual representation (reviewed in Chapter 2). First, we collected the largest dataset of socialness ratings available to date, including norms for over 8,000 individual words, to enable future larger scale investigations into the representation of socially-relevant concepts. Second, we quantified socialness using an inclusive definition to avoid biasing observations towards specific aspects of social experience. The resulting ratings were used to explore how socialness relates to (i) other key lexical and semantic variables (e.g., familiarity, abstractness, valence) and (ii) behavioural performance on lexical-semantic tasks. We found that socially-relevant concepts span the entire concreteness continuum, and include concrete concepts like *people*, and more abstract concepts like *rumour* and *friendly*. We also showed that socialness is positively correlated with affective dimensions (i.e., emotional valence, arousal), but the magnitudes of these statistical relationships were modest at best. Furthermore,

we demonstrated for the first time that socialness explains unique variance in behavioural indices of lexical-semantic processing. Together, these findings suggest that socialness captures a distinct and behaviourally-relevant facet of meaning.

In **Chapter 4** we directly tested a prediction of the social semantics framework that the neural correlates of social cognition and semantic cognition overlap. We focused on one particular component of the framework and hypothesized that social tasks would recruit brain regions that are sensitive to increased semantic control demands. In particular, we expected overlap in the left inferior frontal gyrus (LIFG) and posterior middle temporal gyrus because these regions support the retrieval of weak semantic associations and the selection of task-relevant conceptual information from amongst competing semantic representations (Jefferies, 2013; Lambon Ralph et al., 2017). Using a meta-analytic approach that synthesized the results of 499 functional neuroimaging experiments, we demonstrated that the neural correlates underpinning four types of social phenomena (theory of mind, trait inference, empathy and moral reasoning) reliably include the LIFG. The fact that the LIFG is consistently activated across four social abilities is indicative of a fundamental contribution to social cognition. Furthermore, the likelihood of activation of the LIFG was related to the difficulty of social tasks, consistent with a role in cognitive control.

To better understand how the LIFG contributes to social cognition, **Chapter 5** explored its functional organization by performing a large-scale data-driven and bimodal investigation into its voxel-wise patterns of functional connectivity. The results revealed gradual changes in connectivity along two main spatial dimensions that distinguish between LIFG sub-regions affiliated with different large-scale functional networks. While the dorsal LIFG is coupled with domain-general control networks, there seems to be an increasing bias towards the application of control to conceptual representations in the ventral LIFG via stronger coupling with semantic representation areas (e.g., the ATLs). The results are consistent with the view that the LIFG supports multiple forms of cognitive control (e.g., task-general, semantic), which are flexibly recruited in the service of social cognition depending on task demands.

In the next sections, I will discuss the key implications of the thesis with respect to two major research questions outlined in the introductory chapters: 1) are social concepts a unique or special form of conceptual knowledge? and 2) does the semantic control system contribute to social cognition? Then, I will consider broader implications for the development of future models of social cognition.

2. Insights into the representation of socially relevant conceptual knowledge

Multidimensional theories of semantic representation propose that concepts are represented in many ways, including via language, perceptual-motor and emotion systems, amongst others (Borghi et al., 2019; Lambon Ralph et al., 2017; Reilly et al., 2016). It has been proposed that socialness, a term referring to those aspects of meaning derived through our experiences with the social environment, could contribute uniquely to one of these underlying semantic dimensions (Barsalou, 2020; Borghi et al., 2019). However, behavioural evidence in favour of a unique contribution of socialness to conceptual representation is lacking. **Chapter 3** reports the first evidence that the socialness of word meanings can account for variance in several behavioural indices of lexical-semantic processing. Notably, this is unique variance that cannot be explained by basic lexical properties (e.g., letter length, age of acquisition) or other established semantic dimensions that tap into sensorimotor (i.e., concreteness), affective (i.e., valence), and linguistic experience (i.e., semantic diversity). Moreover, the association between socialness and lexical task performance was facilitatory in nature, consistent with a semantic richness effect wherein words associated with relatively more semantic information (e.g., semantic features, associates, sensorimotor information) are processed more efficiently (Pexman, 2012; Pexman et al., 2002). Therefore, **these findings endorse socialness as a distinct facet of meaning** that might enrich conceptual representations.

Our findings suggest that there is something special about socially-relevant concepts, but the neural mechanism by which socialness is attributed to conceptual representations remains unclear. One proposal argues that social concepts are a distinct semantic category represented in a dedicated semantic store, which might be located in the dorso-polar ATL (Olson et al., 2013). Consistent with this, fMRI studies have found that this brain region responds selectively to concepts of a social nature when compared to non-social concepts (e.g., (Binney et al., 2016; Lin et al., 2019; Olson et al., 2013; Simmons et al., 2010; Zahn et al., 2007) and is functionally coupled with other regions often implicated in social cognition, including the right temporo-parietal junction (Simmons et al., 2010). However, there are other parts of the ATL which respond equally to social and non-social concepts (Binney et al., 2016; Rice et al., 2018), and this goes against the proposal that they rely on separate semantic stores. Another possibility, therefore, is that the semantic information captured by the socialness dimension is just one type of input into a system for the representation of all concepts, much like sensorimotor information is (Binney et al., 2016). It has been suggested that the assimilation of such modality-/domain- specific information into coherent conceptual

representations manifests as gradual variation in the functional specialization of ATL sub-regions, which mirrors the underlying connectivity to the 'spokes', and converges into a domain-general modality-invariant centre-point in its ventrolateral portion (Binney et al., 2012; Rice, Hoffman, et al., 2015; also see Plaut (2002) for a computational model). This 'graded hub' hypothesis might explain why social and non-social concepts are associated with both shared and differential activation within the ATLs (Binney et al., 2016; Rice et al., 2018; Rice, Lambon Ralph, et al., 2015).

Our findings provide some insights into the nature of the information captured by the socialness construct and they are seemingly in line with the latter proposal. It has been suggested that socialness is strongly associated with abstractness (Barsalou & Wiemer-Hastings, 2005; Borghi et al., 2019), and affective information (Troche et al., 2014), and also that these relationships might explain at least some of the preferential activation of the dorso-polar ATL to social concepts (e.g., Binney et al., 2016; for a related discussion, see Wang et al., 2019). However, our initial investigations revealed only weak associations between socialness and these established dimensions of meaning, perhaps suggesting that **socialness is a distinct dimension of a singular semantic space,** as has been previously suggested (Borghi et al., 2019). Our findings also imply that social concepts are not a homogeneous semantic category; rather, **there might be sub-types of social concepts** that are differentially related to other conceptual dimensions proposed to define this unitary semantic space. For example, it appears there could be groups of social concepts that reflect (i) the concrete-abstract distinction (e.g., *people* vs *politics*) and (ii) the emotional-neutral distinction (e.g., *friendly* vs *nationwide*). At the neural level, this might manifest as stronger engagement of perceptual-motor cortices by concrete social concepts. In the case of emotional social concepts, we might expect the recruitment of brain areas involved in processing affective information, such as the amygdala (Kuhnke et al., 2022). Indeed, a meta-analysis of fMRI studies has found that social concepts activate the amygdala more consistently than non-social concepts (Rice, Lambon Ralph, et al., 2015). However, we have only begun the scratch the surface of this exciting research area, and future studies are needed to elucidate how socialness fits within a multidimensional semantic space, and how it is represented by the brain.

3. Insights into the cognitive control of social cognition

There is a growing acceptance of a role of a domain-general executive control system, named the MDN, in various aspects of social information processing (e.g., Amodio, 2014; Darda & Ramsey, 2019; Meyer & Lieberman, 2012). The work presented in this thesis addressed

outstanding questions about social regulation, and specifically whether control systems other than the MDN have a role to play. Moreover, we sought to understand whether there are particular task demands that drive their engagement.

Our results are consistent with the view that **social regulation relies partly on semantic control mechanisms** that enable the context- and task- appropriate use of conceptual knowledge. This is supported by the finding that various social tasks reliably activate brain regions that respond to an increased need for semantic control. These control mechanisms are thought to include (i) semantic retrieval processes engaged when weak associations or noncanonical semantic information are task-relevant, and (ii) selection processes that resolve competition between alternative semantic representations (Badre et al., 2005; Bedny et al., 2008; Gold et al., 2006; Snyder et al., 2011). Consistent with this, the LIFG has been shown to respond to semantic control manipulations that probe social concepts (Satpute et al., 2014). However, whether the activation of SCN regions during high-level social tasks also reflects semantic control has yet to be demonstrated directly.

Given that the semantic control network is partly dissociable from the MDN (Gao et al., 2021, 2021; Whitney et al., 2012), our results suggest that **social cognition is regulated by at least two systems for control** – one that is task-general and one that is tuned towards semantic processing. In contrast to the SCN, the MDN is thought to support task-general executive abilities, such as the controlled application of task rules, working memory and response inhibition (Assem et al., 2022; Duncan, 2010). Our findings also suggest that the **LIFG acts as an interface between domain-general control and semantic systems**. This could involve tying together the MDN's role in cognitively challenging processing and the SCN's role in processing meaning, resulting in an intermediate specialization for performing demanding operations specifically on conceptual representations (Chiou et al., 2022). As such, the topographical organisation of the LIFG is well-suited to allow top-down constraints (e.g., behavioural goals) to be imposed on automatic semantic retrieval (also see Chiou et al., 2022; Davey et al., 2016). This might explain why the LIFG is one of the regions activated by all five types of meaning-imbued social tasks investigated, as well as semantic control manipulations.

Some dual-route theories of social cognition postulate that we engage in controlled social processing (and, consequently, rely on control systems) only under certain circumstances, like those involving a deliberate attempt to make a social judgement (e.g., explicit task instructions) (Evans, 2008; Lieberman, 2007; Satpute & Lieberman, 2006). Contrary to the predictions of these models, we found that the LIFG (and most other activation clusters) is equally likely to contribute to both explicit and implicit social tasks. This finding is

in line with an alternative **single-route cognitive architecture in which cognitive control systems are ubiquitously involved** (i.e., regardless of the implicit or explicit nature of the task), but the extent of their contribution varies depending on situational demands. Indeed, we found that the likelihood of activation of SCN regions during social tasks was influenced by the relative task difficulty. This interpretation mirrors the proposal that flexible use of semantic knowledge involves automatic representational processes that offer a default solution (e.g., via automatic spreading of activation), which is evaluated and may be adjusted by control processes to ensure its suitability given the task/context (Jefferies, 2013; Lambon Ralph et al., 2017; also see default-interventionist dual-route theories; Evans, 2011).

4. Towards a neurocognitive model of controlled social cognition

In this section, I describe some broader implications of the current findings, focusing on how they inform the development of neurocognitive models of social cognition. Binney and Ramsey (2020) proposed that social information processing is underpinned by the same cognitive and brain systems as general semantic cognition. This thesis provides the first direct neuroimaging evidence in support of a role of the semantic control system in social tasks. There is also now mounting evidence from neuropsychology and neuroimaging that the ATL semantic hub contributes to social cognition via a role in the representation of concepts (Binney & Ramsey, 2020; Collins & Olson, 2014; Olson et al., 2013). Altogether, this body of research provides empirical support for the involvement of brain regions underpinning both semantic control and representation in social cognition, as predicted by Binney and Ramsey (2020)'s model. This thesis also offers an explanation for the importance of the LIFG in social and semantic cognition by showing that it functions as a nexus between the brain networks recruited for domain-general control and semantic representation, perhaps mediating the integration of information from these two systems (also see Chiou et al., 2022; Davey et al., 2016). Together with findings from other recent lines of research (e.g., Binney & Ramsey, 2020; Darda et al., 2018; Darda & Ramsey, 2019; Meyer & Lieberman, 2012; Zaki et al., 2010), the current work suggests that generalizable systems contribute to social tasks. Therefore, **future models of social cognition should delineate the contributions of domain-general systems**, including those underpinning executive control and semantic cognition. Moreover, the finding that the LIFG is reliably activated by five social abilities demonstrates that the social semantics framework can provide a unifying explanation applicable to diverse forms of social information processing. Thus, this thesis demonstrates that there is merit to studying social cognition from a general semantics perspective.

However, this thesis also suggests that there is something different about socialness. It might seem that there are two opposing approaches to the study of social cognition: 1) strong versions of generalist frameworks proffer that social information processing can be fully explained by generalizable mechanisms (Binney & Ramsey, 2020; Heyes, 2014), whereas 2) domain-specific views champion the uniqueness of social cognitive mechanisms (e.g., Saxe & Kanwisher, 2004). Although domain-general semantic systems contribute to social tasks, the unique contribution of socialness to conceptual representation suggests that social and semantic cognition cannot be fully equated. This highlights the **need for 'hybrid' models that incorporate both specialized and domain-general neurocognitive components**. Indeed, several researchers have recently proposed that social cognition is likely underpinned by a combination of domain-general and domain-specific mechanisms that have complementary roles (Michael & D'Ausilio, 2015; Ramsey, 2018; Ramsey & Ward, 2020; Spunt & Adolphs, 2017; Zaki et al., 2010). The social semantic framework can accommodate a degree of specialization for social stimuli because the overall cognitive architecture does not depend on the extent to which 'spokes' show domain- or modality- selectivity (Binney & Ramsey, 2020). For example, uniquely social information processed by privileged neural pathways for social inputs, such as the face fusiform area and the extrastriate body area (Kanwisher, 2010), might be integrated into coherent meaningful representations computed via general semantic mechanisms implemented in the ATLs. Consistent with this, a number of studies have reported functional coupling between person perception regions and the ATL semantic hub during social tasks (Greven et al., 2016, 2019; Greven & Ramsey, 2017; Wang et al., 2017). Therefore, investigations into the division of labour between domain-specific and domain-general mechanisms, as well as into how they interact in the service of social cognition are a promising avenue for future research.

In light of this, it is important to consider when incorporating domain-specific mechanisms into models of social cognition is warranted. Researchers have recently argued that a higher bar should be set - domain-specific mechanisms should only be considered when the operation of domain-general mechanisms alone fails to account for empirical observations (Binney & Ramsey, 2020; Ramsey & Ward, 2020). The literature on semantic cognition illustrates a case in which converging multimodal evidence suggests that the operation of task-general control processes implemented in the MDN cannot explain the engagement of the ventral LIFG and posterior middle temporal gyrus in cognitively challenging semantic tasks (Chiou et al., 2022; Gao et al., 2021; Jackson, 2021; Whitney et al., 2012). The current results are in line with a relative specialization of the ventral LIFG for semantic processing, showing

that it might arise from its stronger functional connectivity with semantic representation regions. In a similar vein, future social neuroscientific research should explicitly model and empirically test the contribution of domain-general systems alongside any proposed form of specialized processes (Ramsey & Ward, 2020).

This thesis also emphasises the multi-faceted nature of social abilities. Although this might seem like an obvious observation, many publications leave the misleading impression that this is not the case by engaging in the following practices: (i) implicitly using social terms like *ToM* and *empathy* to refer to singular constructs and without explicitly articulating their chosen conceptualization, (ii) glassing over the various forms they can take (e.g., inferring personality traits vs temporary mental states, cognitive versus affective mental states), and (iii) making reference to unitary brain networks named based on the phenomenon under investigation (e.g., the 'ToM network', the 'social brain') (Schaafsma et al., 2015). However, there is relative consensus that complex **social abilities emerge from the concurrent contributions of multiple processes** (Alcalá-López et al., 2018; Schaafsma et al., 2015; Schurz et al., 2020; Singer, 2006; Warnell & Redcay, 2019; Zaki, 2013), just like general semantic abilities do (Lambon Ralph et al., 2017). In line with this, we found that multiple social abilities activated not only a common set of brain regions, but also regions that showed (i) preferential responses to specific social abilities (e.g., somatosensory cortex and posterior LIFG in empathy for pain), and (ii) different relationships to task characteristics (type of instructional cue, relative difficulty), which is indicative of dissociable contributions. Similarly, we found that individual social concepts can be associated with strikingly different levels of sensorimotor and affective information, implying differential reliance on the respective 'spokes'. Together, these findings are consistent with the view that social cognition draws on brain regions/networks that underpin dissociable cognitive resources in a flexible, task-dependent manner (Alcalá-López et al., 2018). Deconstructing social cognition into its elementary building blocks is a major challenge for cognitive neuroscience (Happé et al., 2017; Schaafsma et al., 2015), but, as this thesis shows, insights from parallel literatures concerned with domain-general systems might provide useful clues.

5. Directions for future research

While this thesis provides empirical support in favour of socialness as a distinct aspect of meaning, many questions remain about the representation of socially-relevant concepts in the brain. For example, it is unclear whether the selective activation of the dorso-polar ATL in response to social concepts is driven by socialness per se or other confounding sources of

information. This is because most extant neuroimaging research on social concepts does not control for some of the semantic dimensions identified as important for conceptual representation in the psycholinguistic literature. Two semantic dimensions – valence and abstractness - are particularly worthy of further scrutiny because they have been found to modulate the activation of the temporo-polar ATL (Binder et al., 2005; Ethofer et al., 2006; Hoffman et al., 2015; Mellem et al., 2016), sometimes in a way that interacts with the effect of socialness (Lin et al., 2018; Mellem et al., 2016). Multivariate pattern analyses of functional neuroimaging data might be particularly useful in disentangling the representation of socialness form these other conceptual dimensions given that they can provide insight into the information encoded in activity patterns across multiple voxels (Haxby et al., 2014). For example, to test whether there are any brain regions that represent socialness independently of valence, a classifier can be trained to distinguish between the activity patterns elicited by social and non-social emotionally-valenced words and tested on its ability to discriminate between social and non-social neutral words. Successful classification would be indicative of valence-*independent* representation of social semantic content. A similar approach can be taken to examine whether social conceptual representations that are independent of sensorimotor features exist. In this case, it will also be important to consider finer-grained perceptual-motor dimensions, like those tapping into specific modalities, because these can make dissociable contributions to semantic representation (Lynott et al., 2020). For instance, auditory information seems to be closely related to socialness (Troche et al., 2017; Villani et al., 2019), perhaps reflecting the fundamental role of verbal communication in social experience (Borghi et al., 2019). This is an important relationship to explore because the dorsal ATL also shows a preference for the auditory modality and for linguistic information, likely because of its proximity and connectivity to the auditory cortex (Murphy et al., 2017; Visser et al., 2012; Visser & Lambon Ralph, 2011).

While this thesis implicates regions recruited by semantic control in social cognition, they cannot directly speak to the nature of their contribution. Future functional neuroimaging research can directly test whether regions of the SCN are sensitive to semantic control manipulations within social tasks. This could be accomplished, for example, by manipulating the strength of the association between a social stimulus, such as a facial emotion portrayal, and its interpretation. If the LIFG and posterior middle temporal gyrus implement controlled semantic retrieval processes during social tasks, a weak association should lead to stronger activation when compared to a strong association. Nevertheless, such evidence would be correlational in nature, so neuropsychological and neuromodulation studies are essential for

determining whether these regions are causally involved in the controlled access to social concepts. If empirical evidence supports these predictions, it would be important to further investigate whether socialness modulates the activation of the SCN by concurrently manipulating the task difficulty and the socialness of the stimuli. This is in light of the current findings and recent proposals (e.g., Ramsey & Ward, 2020) which suggest that 'hybrid' approaches are needed, and these include both testing the contribution of generalizable systems as well as any domain-specific ways in which they might serve social cognition. Previous fMRI studies have found that social concepts (compared to non-social concepts) elicit stronger BOLD responses in semantic control regions (Binney et al., 2016; Mellem et al., 2016), perhaps suggesting that the retrieval of social concepts is more demanding. If this is true, manipulations of task difficulty and stimulus socialness should have an interactive effect on SCN activation, such that the effect of difficulty is stronger for social stimuli. It is also possible that the controlled retrieval of social concepts relies on additional brain regions, much like social concepts selectively recruit the dorso-polar ATLs (Binney et al., 2016), or that it is associated with preferential engagement of SCN sub-regions, like is the case for other stimulus characteristics (e.g., verbal/non-verbal format; Krieger-Redwood et al., 2015).

6. Conclusions

This thesis sought to test whether studying social cognition through a semantic lens can provide insights into its neurocognitive bases. As predicted by this generalist approach, it was found that the neural correlates of social and general semantic cognition overlap, particularly in brain regions that respond to increased demands for cognitive control. Conversely, we also found that there is something special about socialness as an aspect of meaning, which distinguishes constructs like *politics, friendship* and *romance* from those devoid of social significance. These findings have important implications for models of social cognition and advocate for the adoption of 'hybrid' approaches that place similar emphasis on the investigation of both domain-general and domain-specific contributions. While the current findings can only shed so much light on what is a complex picture, they nonetheless demonstrate the utility of investigating social information processing from a general semantics perspective. By doing so, this thesis opens exciting avenues for future social neuroscientific research, and provides scientific resources in the form of openly-available datasets to facilitate such endeavours.

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*. https://doi.org/10.3389/FNINF.2014.00014/BIBTEX

Achim, A. M., Guitton, M., Jackson, P. L., Boutin, A., & Monetta, L. (2013). On what ground do we mentalize? Characteristics of current tasks and sources of information that contribute to mentalizing judgments. *Psychological Assessment*, *25*, 117–126. https://doi.org/10.1037/a0029137

Ackerman, B. P. (1983). Form and function in children's understanding of ironic utterances. *Journal of Experimental Child Psychology*, *35*(3), 487–508. https://doi.org/10.1016/0022-0965(83)90023-1

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience 2003 4:3*, *4*(3), 165–178. https://doi.org/10.1038/nrn1056

Adolphs, R. (2009). The Social Brain: Neural Basis of Social Knowledge. *Annual Review of Psychology*, *60*(1), 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514

Adolphs, R. (2010). Conceptual Challenges and Directions for Social Neuroscience. *Neuron*, *65*(6), 752–767. https://doi.org/10.1016/j.neuron.2010.03.006

Alam, G. T., Murphy, C., Smallwood, J., & Jefferies, E. (2018). Meaningful inhibition: Exploring the role of meaning and modality in response inhibition. *NeuroImage*, *181*, 108–119. https://doi.org/10.1016/j.neuroimage.2018.06.074

Alcalá-López, D., Smallwood, J., Jefferies, E., Van Overwalle, F., Vogeley, K., Mars, R. B., Turetsky, B. I., Laird, A. R., Fox, P. T., Eickhoff, S. B., & Bzdok, D. (2018). Computing the social brain connectome across systems and states. *Cerebral Cortex*, *28*(7), 2207–2232. https://doi.org/10.1093/cercor/bhx121

Alcalá-López, D., Vogeley, K., Binkofski, F., & Bzdok, D. (2019). Building blocks of social cognition: Mirror, mentalize, share? *Cortex*, *118*, 4–18. https://doi.org/10.1016/j.cortex.2018.05.006

Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. In *Nature Reviews Neuroscience* (Vol. 15, Issue 10, pp. 670–682). Nature Publishing Group. https://doi.org/10.1038/nrn3800

Amodio, D. M. (2019). Social Cognition 2.0: An Interactive Memory Systems Account. *Trends in Cognitive Sciences*, *23*(1), 21–33. https://doi.org/10.1016/j.tics.2018.10.002

REFERENCES

Amodio, D. M., & Cikara, M. (2021). The Social Neuroscience of Prejudice. *Annual Review of Psychology*, *72*(1), 439–469. https://doi.org/10.1146/annurev-psych-010419-050928

Amunts, K., Lenzen, M., Friederici, A. D., Schleicher, A., Morosan, P., Palomero-Gallagher, N., & Zilles, K. (2010). Broca's Region: Novel Organizational Principles and Multiple Receptor Mapping. *PLOS Biology*, *8*(9), e1000489. https://doi.org/10.1371/JOURNAL.PBIO.1000489

Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H. B. M., Zilles, K., & Vogt, O. (1999). Broca's Region Revisited: Cytoarchitecture and Intersubject Variability. *J. Comp. Neurol*, *412*, 319–341. https://doi.org/10.1002/(SICI)1096-9861(19990920)412:2

Anwander, A., Tittgemeyer, M., Von Cramon, D. Y., Friederici, A. D., & Knösche, T. R. (2007). Connectivity-Based Parcellation of Broca's Area. *Cerebral Cortex*, *17*(4), 816–825. https://doi.org/10.1093/CERCOR/BHK034

Anzellotti, S., & Caramazza, A. (2016). From Parts to Identity: Invariance and Sensitivity of Face Representations to Different Face Halves. *Cerebral Cortex*, *26*(5), 1900–1909. https://doi.org/10.1093/cercor/bhu337

Anzellotti, S., Fairhall, S. L., & Caramazza, A. (2014). Decoding Representations of Face Identity That are Tolerant to Rotation. *Cerebral Cortex*, *24*(8), 1988–1995. https://doi.org/10.1093/cercor/bht046

Arioli, M., Basso, G., Poggi, P., & Canessa, N. (2021). Fronto-temporal brain activity and connectivity track implicit attention to positive and negative social words in a novel socio-emotional Stroop task. *NeuroImage*, *226*, 117580. https://doi.org/10.1016/j.neuroimage.2020.117580

Arioli, M., Gianelli, C., & Canessa, N. (2021). Neural representation of social concepts: A coordinate-based meta-analysis of fMRI studies. *Brain Imaging and Behavior*, *15*(4), 1912–1921. https://doi.org/10.1007/S11682-020-00384-6

Asano, R., Boeckx, C., & Seifert, U. (2021). Hierarchical control as a shared neurocognitive mechanism for language and music. *Cognition*, *216*, 104847. https://doi.org/10.1016/j.cognition.2021.104847

Assem, M., Glasser, M. F., Van Essen, D. C., & Duncan, J. (2020). A Domain-General Cognitive Core Defined in Multimodally Parcellated Human Cortex. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhaa023

Austen, J. (2019). *Pride and prejudice* (J. Kinsley, Ed.; 3rd ed.). Oxford University Press. (Original work published 1813)

REFERENCES

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception: Research article. *Psychological Science*, *19*(7), 724–732. https://doi.org/10.1111/j.1467-9280.2008.02148.x

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, *12*(5), 193–200. https://doi.org/10.1016/J.TICS.2008.02.004

Badre, D., & D'Esposito, M. (2007). Functional Magnetic Resonance Imaging Evidence for a Hierarchical Organization of the Prefrontal Cortex. *Journal of Cognitive Neuroscience*, *19*(12), 2082–2099. https://doi.org/10.1162/jocn.2007.19.12.2082

Badre, D., & Desrochers, T. M. (2019). Hierarchical cognitive control and the frontal lobes. *Handbook of Clinical Neurology*, *163*, 165–177. https://doi.org/10.1016/B978-0-12-804281-6.00009-4

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, *45*(13), 2883–2901. https://doi.org/10.1016/j.neuropsychologia.2007.06.015

Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, *47*(6), 907–918. https://doi.org/10.1016/j.neuron.2005.07.023

Bahlmann, J., Blumenfeld, R. S., & D'Esposito, M. (2015). The Rostro-Caudal Axis of Frontal Cortex Is Sensitive to the Domain of Stimulus Information. *Cerebral Cortex*, *25*(7), 1815–1826. https://doi.org/10.1093/cercor/bht419

Bailey, P., & Von Bonin, G. (1951). The isocortex of man. In *Illinois Monographs in the Medical Sciences*.

Bajada, C. J., Costa Campos, L. Q., Caspers, S., Muscat, R., Parker, G. J. M., Lambon Ralph, M. A., Cloutman, L. L., & Trujillo-Barreto, N. J. (2020a). A tutorial and tool for exploring feature similarity gradients with MRI data. *NeuroImage*, *221*, 117140. https://doi.org/10.1016/j.neuroimage.2020.117140

Bajada, C. J., Jackson, R. L., Haroon, H. A., Azadbakht, H., Parker, G. J. M., Lambon Ralph, M. A., & Cloutman, L. L. (2017). A graded tractographic parcellation of the temporal lobe. *NeuroImage*, *155*, 503–512. https://doi.org/10.1016/J.NEUROIMAGE.2017.04.016

REFERENCES

Bajada, C. J., Trujillo-Barreto, N. J., Parker, G. J. M., Cloutman, L. L., & Lambon Ralph, M. A. (2019). A structural connectivity convergence zone in the ventral and anterior temporal lobes: Data-driven evidence from structural imaging. *Cortex*, S0010945219302527. https://doi.org/10.1016/j.cortex.2019.06.014

Balgova, E., Diveica, V., Walbrin, J., & Binney, R. J. (2021). The Ventrolateral Anterior Temporal Lobe is Commonly Engaged by Both Mental State Inference and Semantic Association Tasks. *BioRxiv*, 2021.09.10.459496. https://doi.org/10.1101/2021.09.10.459496

Balgova, E., Diveica, V., Walbrin, J., & Binney, R. J. (2022). The role of the ventrolateral anterior temporal lobes in social cognition. *Human Brain Mapping*, *43*(15), 4589-4608. https://doi.org/10.1002/HBM.25976

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The english lexicon project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46. https://doi.org/10.1016/0010-0277(85)90022-8

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241–251. https://doi.org/10.1017/S0021963001006643

Barredo, J., Öztekin, I., & Badre, D. (2015). Ventral Fronto-Temporal Pathway Supporting Cognitive Control of Episodic Memory Retrieval. *Cerebral Cortex*, *25*(4), 1004–1019. https://doi.org/10.1093/cercor/bht291

Barredo, J., Verstynen, T. D., & Badre, D. (2016). Organization of cortico-cortical pathways supporting memory retrieval across subregions of the left ventrolateral prefrontal cortex. *Journal of Neurophysiology*, *116*(3), 920–937. https://doi.org/10.1152/jn.00157.2016

Barrett, H. C. (2012). A hierarchical model of the evolution of human brain specializations. *Proceedings of the National Academy of Sciences*, *109*, 10733–10740. https://doi.org/10.1073/pnas.1201898109

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in Emotion Perception. *Current Directions in Psychological Science*, *20*(5), 286–290. https://doi.org/10.1177/0963721411422522

REFERENCES

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, *59*, 617–645. https://doi.org/10.1146/ANNUREV.PSYCH.59.103006.093639

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1281. https://doi.org/10.1098/RSTB.2008.0319

Barsalou, L. W. (2020). Challenges and Opportunities for Grounding Cognition. *Journal of Cognition*, *3*(1), 1–24. https://doi.org/10.5334/JOC.116

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating Abstract Concepts. In D. Pecher & R. A. Zwaan (Eds.), *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking* (pp. 129–164). Cambridge University Press.

Bas-Hoogendam, J. M., van Steenbergen, H., Kreuk, T., van der Wee, N. J. A., & Westenberg, P. M. (2017). How embarrassing! The behavioral and neural correlates of processing social norm violations. *PLOS ONE*, *12*(4), e0176326. https://doi.org/10.1371/journal.pone.0176326

Bedny, M., McGill, M., & Thompson-Schill, S. L. (2008). Semantic Adaptation and Competition during Word Comprehension. *Cerebral Cortex*, *18*(11), 2574–2585. https://doi.org/10.1093/cercor/bhn018

Betti, V., & Aglioti, S. M. (2016). Dynamic construction of the neural networks underpinning empathy for pain. *Neuroscience & Biobehavioral Reviews*, *63*, 191–206. https://doi.org/10.1016/j.neubiorev.2016.02.009

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*(11), 527–536. https://doi.org/10.1016/j.tics.2011.10.001

Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M., & Desai, R. H. (2016). Toward a brain-based componential semantic representation. *Cognitive Neuropsychology*, *33*(3–4), 130–174. https://doi.org/10.1080/02643294.2016.1147426

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*, *19*, 2767–2796. https://doi.org/10.1093/cercor/bhp055

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct Brain Systems for Processing Concrete and Abstract Concepts. *Journal of Cognitive Neuroscience*, *17*(6), 905–917. https://doi.org/10.1162/0898929054021102

Binney, R. J., & Lambon Ralph, M. A. (2015). Using a combination of fMRI and anterior temporal lobe rTMS to measure intrinsic and induced activation changes across the

semantic cognition network. *Neuropsychologia*, *76*, 170–181. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2014.11.009

Binney, R. J., & Ramsey, R. (2020). Social Semantics: The role of conceptual knowledge and cognitive control in a neurobiological model of the social brain. *Neuroscience & Biobehavioral Reviews*, *112*, 28–38. https://doi.org/10.1016/j.neubiorev.2020.01.030

Binney, R. J., Embleton, K. V, Jefferies, E., Parker, G. J. M., & Lambon Ralph, M. A. (2010). The Ventral and Inferolateral Aspects of the Anterior Temporal Lobe Are Crucial in Semantic Memory: Evidence from a Novel Direct Comparison of Distortion-Corrected fMRI, rTMS, and Semantic Dementia. *Cerebral Cortex*, *20*, 2728–2738. https://doi.org/10.1093/cercor/bhq019

Binney, R. J., Henry, M. L., Babiak, M., Pressman, P. S., Santos-Santos, M. A., Narvid, J., Mandelli, M. L., Strain, P. J., Miller, B. L., Rankin, K. P., Rosen, H. J., & Gorno-Tempini, M. L. (2016). Reading words and other people: A comparison of exception word, familiar face and affect processing in the left and right temporal variants of primary progressive aphasia. *Cortex*, *82*, 147–163. https://doi.org/10.1016/j.cortex.2016.05.014

Binney, R. J., Hoffman, P., & Lambon Ralph, M. A. (2016). Mapping the Multiple Graded Contributions of the Anterior Temporal Lobe Representational Hub to Abstract and Social Concepts: Evidence from Distortion-corrected fMRI. *Cerebral Cortex*, *26*(11), 4227–4241. https://doi.org/10.1093/cercor/bhw260

Binney, R. J., Parker, G. J. M., & Lambon Ralph, M. A. (2012). Convergent connectivity and graded specialization in the Rostral human temporal Lobe as revealed by diffusion-weighted imaging probabilistic tractography. *Journal of Cognitive Neuroscience*, *24*(10), 1998–2014. https://doi.org/10.1162/jocn_a_00263

Binney, R. J., Zuckerman, B., & Reilly, J. (2016). A Neuropsychological Perspective on Abstract Word Representation: From Theory to Treatment of Acquired Language Disorders. *Current Neurology and Neuroscience Reports*, *16*(9), 1–8. https://doi.org/10.1007/S11910-016-0683-0/FIGURES/1

Bonini, L., Rotunno, C., Arcuri, E., & Gallese, V. (2022). Mirror neurons 30 years later: Implications and applications. *Trends in Cognitive Sciences*, *26*(9), 767–781. https://doi.org/10.1016/j.tics.2022.06.003

Borghi, A. M., & Binkofski, F. (2014). *Words as Social Tools: An Embodied View on Abstract Concepts*. Springer. https://doi.org/10.1007/978-1-4614-9539-0

REFERENCES

Borghi, A. M., Barca, L., Binkofski, F., & Tummolini, L. (2018). Varieties of abstract concepts: Development, use and representation in the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752). https://doi.org/10.1098/RSTB.2017.0121

Borghi, A. M., Barca, L., Binkofski, F., Castelfranchi, C., Pezzulo, G., & Tummolini, L. (2019). Words as social tools: Language, sociality and inner grounding in abstract concepts. *Physics of Life Reviews*, *29*, 120–153. https://doi.org/10.1016/J.PLREV.2018.12.001

Bosco, F. M., Tirassa, M., & Gabbatore, I. (2018). Why pragmatics and Theory of Mind do not (Completely) overlap. *Frontiers in Psychology*, *9*, 1453. https://doi.org/10.3389/fpsyg.2018.01453

Bottenhorn, K. L., Salo, T., Sutherland, M. T., & Laird, A. L. (2018, August 8). Quantitative comparison of functional decoding approaches across meta-analytic frameworks. *F1000Research*. Neuroinformatics. https://doi.org/10.7490/f1000research.1115906.1

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J. A., Adcock, R. A., Avesani, P., Baczkowski, B. M., Bajracharya, A., Bakst, L., Ball, S., Barilari, M., Bault, N., Beaton, D., Beitner, J., … Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, *582*(7810), 84–88. https://doi.org/10.1038/s41586-020-2314-9

Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, *38*(9), 1207–1215. https://doi.org/10.1016/S0028-3932(00)00034-8

Branzi, F. M., Humphreys, G. F., Hoffman, P., & Lambon Ralph, M. A. (2020). Revealing the neural networks that extract conceptual gestalts from continuously evolving or changing semantic contexts. *NeuroImage*, *220*(116802). https://doi.org/10.1016/j.neuroimage.2020.116802

Brass, M., & Heyes, C. (2005). Imitation: Is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*(10), 489–495. https://doi.org/10.1016/j.tics.2005.08.007

Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2359–2367. https://doi.org/10.1098/rstb.2009.0066

Braunsdorf, M., Blazquez Freches, G., Roumazeilles, L., Eichert, N., Schurz, M., Uithol, S., Bryant, K. L., & Mars, R. B. (2021). Does the temporal cortex make us human? A

review of structural and functional diversity of the primate temporal lobe. *Neuroscience & Biobehavioral Reviews*, *131*, 400–410. https://doi.org/10.1016/J.NEUBIOREV.2021.08.032

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig : Barth.

Brooks, J. A., & Freeman, J. B. (2018). Conceptual knowledge predicts the representational structure of facial emotion perception. *Nature Human Behaviour*, *2*(8), Article 8. https://doi.org/10.1038/s41562-018-0376-6

Brothers, L. (1990). The social brain: A project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience*, *1*, 27–51.

Brown, S., & Schafer, E. A. (1888). An investigation into the functions of the occipital and temporal lobes of the monkey's brain. *Philosophical Transactions of the Royal Society of London. (B.)*, *179*, 303–327. https://doi.org/10.1098/rstb.1888.0011

Brysbaert, M. (2019). How Many Participants Do We Have to Include in Properly Powered Experiments? A Tutorial of Power Analysis with Reference Tables. *Journal of Cognition*, *2*(1), 1–38. https://doi.org/10.5334/joc.72

Brysbaert, M., & Biemiller, A. (2017). Test-based age-of-acquisition norms for 44 thousand English word meanings. *Behavior Research Methods*, *49*(4), 1520–1523. https://doi.org/10.3758/s13428-016-0811-4

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., Mandera, P., McCormick, S. F., & Keuleers, E. (2018). Word prevalence norms for 62,000 English lemmas. *Behavior Research Methods 2018 51:2*, *51*(2), 467–479. https://doi.org/10.3758/S13428-018-1077-9

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods 2012 44:4*, *44*(4), 991–997. https://doi.org/10.3758/S13428-012-0190-4

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

REFERENCES

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. https://doi.org/10.1196/annals.1440.011

Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., Andrews-Hanna, J. R., Sperling, R. A., & Johnson, K. A. (2009). Cortical Hubs Revealed by Intrinsic Functional Connectivity: Mapping, Assessment of Stability, and Relation to Alzheimer's Disease. *Journal of Neuroscience*, *29*(6), 1860–1873. https://doi.org/10.1523/JNEUROSCI.5062-08.2009

Buckner, R. L., Snyder, A. Z., Shannon, B. J., LaRossa, G., Sachs, R., Fotenos, A. F., Sheline, Y. I., Klunk, W. E., Mathis, C. A., Morris, J. C., & Mintun, M. A. (2005). Molecular, Structural, and Functional Characterization of Alzheimer's Disease: Evidence for a Relationship between Default Activity, Amyloid, and Memory. *Journal of Neuroscience*, *25*(34), 7709–7717. https://doi.org/10.1523/JNEUROSCI.2177-05.2005

Bukowski, H. (2018). The Neural Correlates of Visual Perspective Taking: A Critical Review. *Current Behavioral Neuroscience Reports*, *5*(3), 189–197. https://doi.org/10.1007/s40473-018-0157-6

Burgess, G. C., Kandala, S., Nolan, D., Laumann, T. O., Power, J. D., Adeyemo, B., Harms, M. P., Petersen, S. E., & Barch, D. M. (2016). Evaluation of Denoising Strategies to Address Motion-Correlated Artifacts in Resting-State Functional Magnetic Resonance Imaging Data from the Human Connectome Project. *Https://Home.Liebertpub.Com/Brain*, *6*(9), 669–680. https://doi.org/10.1089/BRAIN.2016.0435

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The Human Hippocampus and Spatial and Episodic Memory. *Neuron*, *35*(4), 625–641. https://doi.org/10.1016/S0896-6273(02)00830-9

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Bzdok, D., Heeger, A., Langner, R., Laird, A. R., Fox, P. T., Palomero-Gallagher, N., Vogt, B. A., Zilles, K., & Eickhoff, S. B. (2015). Subspecialization in the human posterior medial cortex. *NeuroImage*, *106*, 55–71. https://doi.org/10.1016/J.NEUROIMAGE.2014.11.009

REFERENCES

Bzdok, D., Langner, R., Caspers, S., Kurth, F., Habel, U., Zilles, K., Laird, A., & Eickhoff, S. B. (2011). ALE meta-analysis on facial judgments of trustworthiness and attractiveness. *Brain Structure and Function*, *215*(3–4), 209–223. https://doi.org/10.1007/s00429-010-0287-4

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, *217*(4), 783–796. https://doi.org/10.1007/s00429-012-0380-y

Cabeza, R., Ciaramelli, E., & Moscovitch, M. (2012). Cognitive contributions of the ventral parietal cortex: An integrative theoretical account. *Trends in Cognitive Sciences*, *16*(6), 338–352. https://doi.org/10.1016/j.tics.2012.04.008

Camilleri, J. A., Müller, V. I., Fox, P., Laird, A. R., Hoffstaedter, F., Kalenscher, T., & Eickhoff, S. B. (2018). Definition and characterization of an extended multiple-demand network. *NeuroImage*, *165*, 138–147. https://doi.org/10.1016/j.neuroimage.2017.10.020

Carp, J. (2012). On the Plurality of (Methodological) Worlds: Estimating the Analytic Flexibility of fMRI Experiments. *Frontiers in Neuroscience*, *6*, 149. https://doi.org/10.3389/fnins.2012.00149

Carruthers, P. (2016). Two Systems for Mindreading? *Review of Philosophy and Psychology*, *7*(1), 141–162. https://doi.org/10.1007/s13164-015-0259-y

Carruthers, P. (2017). Mindreading in adults: Evaluating two-systems views. *Synthese*, *194*(3), 673–688. https://doi.org/10.1007/s11229-015-0792-3

Carter, R. M., & Huettel, S. A. (2013). A nexus model of the temporal–parietal junction. *Trends in Cognitive Sciences*, *17*(7), 328–336. https://doi.org/10.1016/j.tics.2013.05.007

Cassidy, B. S., & Gutchess, A. H. (2015). Neural responses to appearance-behavior congruity. *Social Cognition*, *33*(3), 211–226. https://doi.org/10.1521/soco.2015.33.3.1

Castelli, F., Frith, C., Happe, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, *125*(8), 1839–1849. https://doi.org/10.1093/brain/awf189

Catricalà, E., Conca, F., Borsa, V. M., Cotelli, M., Manenti, R., Gobbi, E., Binetti, G., Cotta Ramusino, M., Perini, G., Costa, A., Rusconi, M. L., & Cappa, S. F. (2021). Different types of abstract concepts: Evidence from two neurodegenerative patients. *Neurocase*. https://doi.org/10.1080/13554794.2021.1931345

REFERENCES

Catricalà, E., Conca, F., Fertonani, A., Miniussi, C., & Cappa, S. F. (2020). State-dependent TMS reveals the differential contribution of ATL and IPS to the representation of abstract concepts related to social and quantity knowledge. *Cortex*, *123*, 30–41. https://doi.org/10.1016/J.CORTEX.2019.09.018

Catricalà, E., Della Rosa, P. A., Plebani, V., Vigliocco, G., & Cappa, S. F. (2014). Abstract and concrete categories? Evidences from neurodegenerative diseases. *Neuropsychologia*, *64*, 271–281. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2014.09.041

Cerliani, L., Thomas, R. M., Jbabdi, S., Siero, J. C. W., Nanetti, L., Crippa, A., Gazzola, V., D'Arceuil, H., & Keysers, C. (2012). Probabilistic tractography recovers a rostrocaudal trajectory of connectivity variability in the human insular cortex. *Human Brain Mapping*, *33*(9), 2005–2034. https://doi.org/10.1002/hbm.21338

Chan, A. M., Baker, J. M., Eskandar, E., Schomer, D., Ulbert, I., Marinkovic, K., Cash, S. S., & Halgren, E. (2011). First-Pass Selectivity for Semantic Categories in Human Anteroventral Temporal Lobe. *Journal of Neuroscience*, *31*(49), 18119–18129. https://doi.org/10.1523/JNEUROSCI.3122-11.2011

Chan, D., Anderson, V., Pijnenburg, Y., Whitwell, J., Barnes, J., Scahill, R., Stevens, J. M., Barkhof, F., Scheltens, P., Rossor, M. N., & Fox, N. C. (2009). The clinical profile of right temporal lobe atrophy. *Brain*, *132*, 1287–1298. https://doi.org/10.1093/brain/awp037

Chang, C.-F., Hsu, T.-Y., Tseng, P., Liang, W.-K., Tzeng, O. J. L., Hung, D. L., & Juan, C.-H. (2013). Right temporoparietal junction and attentional reorienting. *Human Brain Mapping*, *34*(4), 869–877. https://doi.org/10.1002/hbm.21476

Chapman, C. A., Hasan, O., Schulz, P. E., & Martin, R. C. (2020). Evaluating the distinction between semantic knowledge and semantic access: Evidence from semantic dementia and comprehension-impaired stroke aphasia. *Psychonomic Bulletin and Review*, *27*(4), 607–639. https://doi.org/10.3758/S13423-019-01706-6/FIGURES/2

Chapman, E., Baron-Cohen, S., Auyeung, B., Knickmeyer, R., Taylor, K., & Hackett, G. (2006). Fetal testosterone and empathy: Evidence from the Empathy Quotient (EQ) and the "Reading the Mind in the Eyes" Test. *Social Neuroscience*, *1*(2), 135–148. https://doi.org/10.1080/17470910600992239

Chen, A. C., Welsh, R. C., Liberzon, I., & Taylor, S. F. (2010). 'Do I like this person?' A network analysis of midline cortex during a social preference task. *NeuroImage*, *51*(2), 930–939. https://doi.org/10.1016/j.neuroimage.2010.02.044

REFERENCES

Chen, Y., Shimotake, A., Matsumoto, R., Kunieda, T., Kikuchi, T., Miyamoto, S., Fukuyama, H., Takahashi, R., Ikeda, A., & Lambon Ralph, M. A. (2016). The 'when' and 'where' of semantic coding in the anterior temporal lobe: Temporal representational similarity analysis of electrocorticogram data. *Cortex*, *79*, 1–13. https://doi.org/10.1016/j.cortex.2016.02.015

Chiou, R., & Lambon Ralph, M. A. (2019). Unveiling the dynamic interplay between the hub- and spoke-components of the brain's semantic system and its impact on human behaviour. *NeuroImage*, *199*, 114–126. https://doi.org/10.1016/J.NEUROIMAGE.2019.05.059

Chiou, R., Humphreys, G. F., Jung, J., & Lambon Ralph, M. A. (2018). Controlled semantic cognition relies upon dynamic and flexible interactions between the executive 'semantic control' and hub-and-spoke 'semantic representation' systems. *Cortex*, *103*, 100–116. https://doi.org/10.1016/j.cortex.2018.02.018

Chiou, R., Jefferies, E., Duncan, J., Humphreys, G. F., & Lambon Ralph, M. A. (2022). A middle ground where executive control meets semantics: The neural substrates of semantic control are topographically sandwiched between the multiple-demand and default-mode systems. *Cerebral Cortex*, *bhac358*. https://doi.org/10.1093/cercor/bhac358

Cho, J. W., Korchmaros, A., Vogelstein, J. T., Milham, M. P., & Xu, T. (2021). Impact of concatenating fMRI data on reliability for functional connectomics. *NeuroImage*, *226*, 117549. https://doi.org/10.1016/J.NEUROIMAGE.2020.117549

Chong, T. T.-J., Cunnington, R., Williams, M. A., Kanwisher, N., & Mattingley, J. B. (2008). FMRI Adaptation Reveals Mirror Neurons in Human Inferior Parietal Cortex. *Current Biology*, *18*(20), 1576–1580. https://doi.org/10.1016/j.cub.2008.08.068

Chou, Y. H., Sundman, M., Whitson, H. E., Gaur, P., Chu, M. L., Weingarten, C. P., Madden, D. J., Wang, L., Kirste, I., Joliot, M., Diaz, M. T., Li, Y. J., Song, A. W., & Chen, N. K. (2017). Maintenance and Representation of Mind Wandering during Resting-State fMRI. *Scientific Reports 2017 7:1*, *7*(1), 1–11. https://doi.org/10.1038/srep40722

Clos, M., Amunts, K., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2013). Tackling the multifunctional nature of Broca's region meta-analytically: Co-activation-based parcellation of area 44. *NeuroImage*, *83*, 174–188. https://doi.org/10.1016/J.NEUROIMAGE.2013.06.041

REFERENCES

Cloutier, J., Gabrieli, J. D. E., O'Young, D., & Ambady, N. (2011). An fMRI study of violations of social expectations: When people are not who we expect them to be. *NeuroImage*, *57*(2), 583–588. https://doi.org/10.1016/j.neuroimage.2011.04.051

Cloutman, L. L., & Lambon Ralph, M. A. (2012). Connectivity-based structural and functional parcellation of the human cortex using diffusion imaging and tractography. *Frontiers in Neuroanatomy*, *0*(AUG 2012), 1–18. https://doi.org/10.3389/FNANA.2012.00034/BIBTEX

Coccia, M., Bartolini, M., Luzzi, S., Provinciali, L., & Lambon Ralph, M. A. (2004). Cognitive Neuropsychology Semantic memory is an amodal, dynamic system: Evidence from the interaction of naming and object use in semantic dementia SEMANTIC MEMORY IS AN AMODAL, DYNAMIC SYSTEM: EVIDENCE FROM THE INTERACTION OF NAMING AND OBJECT USE IN SEM. *Cognitive Neuropsychology*, *21*(5), 513–527. https://doi.org/10.1080/02643290342000113

Coifman, R. R., & Lafon, S. (2006). Diffusion maps. *Applied and Computational Harmonic Analysis*, *21*(1), 5–30. https://doi.org/10.1016/J.ACHA.2006.04.006

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., & Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience 2013 16:9*, *16*(9), 1348–1355. https://doi.org/10.1038/nn.3470

Collins, J. A., & Olson, I. R. (2014). Beyond the FFA: The role of the ventral anterior temporal lobes in face processing. *Neuropsychologia*, *61*(1), 65–79. https://doi.org/10.1016/j.neuropsychologia.2014.06.005

Conca, F., Borsa, V. M., Cappa, S. F., & Catricalà, E. (2021). The multidimensionality of abstract concepts: A systematic review. *Neuroscience & Biobehavioral Reviews*, *127*, 474–491. https://doi.org/10.1016/J.NEUBIOREV.2021.05.004

Corbett, F., Jefferies, E., & Lambon Ralph, M. A. (2011). Deregulated Semantic Cognition Follows Prefrontal and Temporo-parietal Damage: Evidence from the Impact of Task Constraint on Nonverbal Object Use. *Journal of Cognitive Neuroscience*, *23*(5), 1125–1135. https://doi.org/10.1162/jocn.2010.21539

Corbett, F., Jefferies, E., Burns, A., & Lambon Ralph, M. A. (2015). Deregulated semantic cognition contributes to object-use deficits in Alzheimer's disease: A comparison with semantic aphasia and semantic dementia. *Journal of Neuropsychology*, *9*(2), 219–241. https://doi.org/10.1111/jnp.12047

REFERENCES

Corbett, F., Jefferies, E., Ehsan, S., & Lambon Ralph, M. A. (2009). Different impairments of semantic cognition in semantic dementia and semantic aphasia: Evidence from the non-verbal domain. *Brain*, *132*(9), 2593–2608. https://doi.org/10.1093/brain/awp146

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), Article 3. https://doi.org/10.1038/nrn755

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers 2004 36:3*, *36*(3), 384–387. https://doi.org/10.3758/BF03195585

Coutanche, M. N., & Thompson-Schill, S. L. (2015). Creating Concepts from Converging Features in Human Cortex. *Cerebral Cortex*, *25*(9), 2584–2593. https://doi.org/10.1093/cercor/bhu057

Crittenden, B. M., Mitchell, D. J., & Duncan, J. (2015). Recruitment of the default mode network during a demanding act of executive control. *ELife*, *4*, e06481. https://doi.org/10.7554/eLife.06481

Cross, K. A., Torrisi, S., Reynolds Losin, E. A., & Iacoboni, M. (2013). Controlling automatic imitative tendencies: Interactions between mirror neuron and cognitive control systems. *NeuroImage*, *83*, 493–504. https://doi.org/10.1016/j.neuroimage.2013.06.060

Crutch, S. J., & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*(3), 615–627. https://doi.org/10.1093/BRAIN/AWH349

Crutch, S. J., Ridha, B. H., & Warrington, E. K. (2007). *The Different Frameworks Underlying Abstract and Concrete Knowledge: Evidence from a Bilingual Patient with a Semantic Refractory Access Dysphasia*. *12*(3), 151–163. https://doi.org/10.1080/13554790600598832

Crutch, S. J., Williams, P., Ridgway, G. R., & Borgenicht, L. (2012). The role of polarity in antonym and synonym conceptual knowledge: Evidence from stroke aphasia and multidimensional ratings of abstract words. *Neuropsychologia*, *50*(11), 2636–2644. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2012.07.015

Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A Review of the Concept. *Emotion Review*, *8*(2), 144–153. https://doi.org/10.1177/1754073914558466

REFERENCES

Cumming, G. (2014). The New Statistics: Why and How. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *Trends in Cognitive Sciences*, *11*(3), 97–104. https://doi.org/10.1016/j.tics.2006.12.005

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, *66*, 4–19. https://doi.org/10.1016/j.jesp.2015.07.006

Dale, E., & O'Rourke, J. (1981). *The Living Word Vocabulary, the Words We Know: A National Vocabulary Inventory*. World Book.

Damoiseaux, J. S., & Greicius, M. D. (2009). Greater than the sum of its parts: A review of studies combining structural connectivity and resting-state functional connectivity. *Brain Structure and Function*, *213*(6), 525–533. https://doi.org/10.1007/S00429-009-0208-6/FIGURES/2

Darda, K. M., & Ramsey, R. (2019). The inhibition of automatic imitation: A meta-analysis and synthesis of fMRI studies. *NeuroImage*, *197*, 320–329. https://doi.org/10.1016/j.neuroimage.2019.04.059

Darda, K. M., Butler, E. E., & Ramsey, R. (2018). Functional Specificity and Sex Differences in the Neural Circuits Supporting the Inhibition of Automatic Imitation. *Journal of Cognitive Neuroscience*, *30*(6), 914–933. https://doi.org/10.1162/jocn_a_01261

Davey, J., Cornelissen, P. L., Thompson, H. E., Sonkusare, S., Hallam, G., Smallwood, J., & Jefferies, E. (2015). Automatic and controlled semantic retrieval: TMS reveals distinct contributions of posterior middle temporal gyrus and angular gyrus. *Journal of Neuroscience*, *35*(46), 15230–15239. https://doi.org/10.1523/JNEUROSCI.4705-14.2015

Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., Krieger-Redwood, K., Bernhardt, B. C., Smallwood, J., & Jefferies, E. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *NeuroImage*, *137*, 165–177. https://doi.org/10.1016/j.neuroimage.2016.05.051

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior Research Methods*, *51*(3), 987–1006. https://doi.org/10.3758/s13428-018-1115-7

REFERENCES

De Greck, M., Scheidt, L., Bölter, A. F., Frommer, J., Ulrich, C., Stockum, E., Enzi, B., Tempelmann, C., Hoffmann, T., Han, S., & Northoff, G. (2012). Altered brain activity during emotional empathy in somatoform disorder. *Human Brain Mapping*, *33*(11), 2666–2685. https://doi.org/10.1002/hbm.21392

De Guzman, M., Bird, G., Banissy, M. J., & Catmur, C. (2016). Self–other control processes in social cognition: From imitation to empathy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150079. https://doi.org/10.1098/rstb.2015.0079

De Lange, F. P., Spronk, M., Willems, R. M., Toni, I., & Bekkering, H. (2008). Complementary Systems for Understanding Action Intentions. *Current Biology*, *18*(6), 454–457. https://doi.org/10.1016/j.cub.2008.02.057

Decety, J., & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*(1), 4–14. https://doi.org/10.1016/j.brainres.2005.12.115

Decety, J., & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. *Neuroscientist*, *13*(6), 580–593. https://doi.org/10.1177/1073858407304654

Deen, B., Koldewyn, K., Kanwisher, N., & Saxe, R. (2015). Functional Organization of Social Perception and Cognition in the Superior Temporal Sulcus. *Cerebral Cortex*, *25*(11), 4596–4609. https://doi.org/10.1093/cercor/bhv111

Dejerine, J., & Dejerine-Klumpke, A. (1895). *Anatomie des centres nerveux*. Rueff.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18. https://doi.org/10.1037/0022-3514.56.1.5

Devlin, J. T., Matthews, P. M., & Rushworth, M. F. S. (2003). Semantic Processing in the Left Inferior Prefrontal Cortex: A Combined Functional Magnetic Resonance Imaging and Transcranial Magnetic Stimulation Study. *Journal of Cognitive Neuroscience*, *15*(1), 71–84. https://doi.org/10.1162/089892903321107837

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Moss, H. E., Matthews, P. M., & Tyler, L. K. (2000). Susceptibility-induced loss of signal: Comparing PET and fMRI on a semantic task. *NeuroImage*, *11*(6 I), 589–600. https://doi.org/10.1006/nimg.2000.0595

REFERENCES

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*(1), 176–180. https://doi.org/10.1007/BF00230027

Diachek, E., Blank, I., Siegelman, M., Affourtit, J., & Fedorenko, E. (2020). The domain-general multiple demand (MD) network does not support core aspects of language comprehension: A large-scale fMRI investigation. *The Journal of Neuroscience*, JN-RM-2036-19. https://doi.org/10.1523/JNEUROSCI.2036-19.2020

Diamond, A. (2013, January 2). Executive functions. In *Annual Review of Psychology*. Annual Reviews Inc. https://doi.org/10.1146/annurev-psych-113011-143750

Dickerson, B. C., & Eichenbaum, H. (2010). The Episodic Memory System: Neurocircuitry and Disorders. *Neuropsychopharmacology*, *35*(1), Article 1. https://doi.org/10.1038/npp.2009.126

Ding, J., Chen, K., Liu, H., Huang, L., Chen, Y., Lv, Y., Yang, Q., Guo, Q., Han, Z., & Lambon Ralph, M. A. (2020). A unified neurocognitive model of semantics language social behaviour and face recognition in semantic dementia. *Nature Communications*, *11*, 2595. https://doi.org/10.1038/s41467-020-16089-9

Diveica, V., Koldewyn, K., & Binney, R. J. (2021). Establishing a role of the semantic control network in social cognitive processing: A meta-analysis of functional neuroimaging studies. *NeuroImage*, *245*, 118702. https://doi.org/10.1016/J.NEUROIMAGE.2021.118702

Diveica, V., Pexman, P. M., & Binney, R. J. (2022). Quantifying social semantics: An inclusive definition of socialness and ratings for 8388 English words. *Behavior Research Methods*, *1*, 1–13. https://doi.org/10.3758/S13428-022-01810-X

Dobbins, I. G., & Wagner, A. D. (2005). Domain-general and Domain-sensitive Prefrontal Mechanisms for Recollecting Events and Detecting Novelty. *Cerebral Cortex*, *15*(11), 1768–1778. https://doi.org/10.1093/CERCOR/BHI054

Dockès, J., Poldrack, R. A., Primet, R., Gözükan, H., Yarkoni, T., Suchanek, F., Thirion, B., & Varoquaux, G. (2020a). Neuroquery, comprehensive meta-analysis of human brain mapping. *ELife*, *9*. https://doi.org/10.7554/ELIFE.53385

Doucet, G., Naveau, M., Petit, L., Zago, L., Crivello, F., Jobard, G., Delcroix, N., Mellet, E., Tzourio-Mazoyer, N., Mazoyer, B., & Joliot, M. (2012). Patterns of hemodynamic low-frequency oscillations in the brain are modulated by the nature of free thought during rest. *NeuroImage*, *59*(4), 3194–3200. https://doi.org/10.1016/J.NEUROIMAGE.2011.11.059

Dove, G. (2016). Three symbol ungrounding problems: Abstract concepts and the future of embodied cognition. *Psychonomic Bulletin & Review*, *23*(4), 1109–1121. https://doi.org/10.3758/s13423-015-0825-4

Dove, G. (2018). Language as a disruptive technology: Abstract concepts, embodiment and the flexible mind. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752). https://doi.org/10.1098/RSTB.2017.0135

Downing, P. E., Chan, A. W.-Y., Peelen, M. V., Dodds, C. M., & Kanwisher, N. (2006). Domain Specificity in Visual Cortex. *Cerebral Cortex*, *16*(10), 1453–1461. https://doi.org/10.1093/cercor/bhj086

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A Cortical Area Selective for Visual Processing of the Human Body. *Science*, *293*(5539), 2470–2473. https://doi.org/10.1126/SCIENCE.1063414

Dricu, M., & Frühholz, S. (2016). Perceiving emotional expressions in others: Activation likelihood estimation meta-analyses of explicit evaluation, passive perception and incidental perception of emotions. *Neuroscience and Biobehavioral Reviews*, *71*, 810–828. https://doi.org/10.1016/j.neubiorev.2016.10.020

Druzgal, T. J., & D'Esposito, M. (2003). Dissecting Contributions of Prefrontal Cortex and Fusiform Face Area to Face Working Memory. *Journal of Cognitive Neuroscience*, *15*(6), 771–784. https://doi.org/10.1162/089892903322370708

Dunbar, R. I. M. (1998). The Social Brain Hypothesis. *Evolutionary Anthropology*, *6*(5), 178–190. https://doi.org/10.1002/(SICI)1520-6505(1998)6:5

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

Duncan, J. (2013). The Structure of Cognition: Attentional Episodes in Mind and Brain. In *Neuron* (Vol. 80, Issue 1, pp. 35–50). Cell Press. https://doi.org/10.1016/j.neuron.2013.09.015

Dunlap, W. P., Jones, M. B., & Bittner, A. C. (2013). Average correlations vs. Correlated averages. *Bulletin of the Psychonomic Society 1983 21:3*, *21*(3), 213–216. https://doi.org/10.3758/BF03334690

Dupuis, M., Meier, E., & Cuneo, F. (2019). Detecting computer-generated random responding in questionnaire-based data: A comparison of seven indices. *Behavior Research Methods*, *51*(5), 2228–2237. https://doi.org/10.3758/s13428-018-1103-y

REFERENCES

Duval, C. L., Bejanin, A., Piolino, P., Laisney, M., De, V., Sayette, L., Belliard, S., Eustache, F., & Atrice Desgranges, B. (2012). Theory of mind impairments in patients with semantic dementia. *Brain*, *135*(1), 228–241. https://doi.org/10.1093/brain/awr309

Eickhoff, S. B., & Grefkes, C. (2011). Approaches for the integrated analysis of structure, function and connectivity of the human brain. *Clinical EEG and Neuroscience*, *42*(2), 107–121. https://doi.org/10.1177/155005941104200211

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage*, *59*(3), 2349–2361. https://doi.org/10.1016/j.neuroimage.2011.09.017

Eickhoff, S. B., Bzdok, D., Laird, A. R., Roski, C., Caspers, S., Zilles, K., & Fox, P. T. (2011). Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation. *NeuroImage*, *57*(3), 938–949. https://doi.org/10.1016/j.neuroimage.2011.05.021

Eickhoff, S. B., Laird, A. R., Fox, P. M., Lancaster, J. L., & Fox, P. T. (2017). Implementation errors in the GingerALE Software: Description and recommendations. *Human Brain Mapping*, *38*(1), 7–11. https://doi.org/10.1002/hbm.23342

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping*, *30*(9), 2907–2926. https://doi.org/10.1002/hbm.20718

Eickhoff, S. B., Nichols, T. E., Laird, A. R., Hoffstaedter, F., Amunts, K., Fox, P. T., Bzdok, D., & Eickhoff, C. R. (2016). Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. *NeuroImage*, *137*, 70–85. https://doi.org/10.1016/j.neuroimage.2016.04.072

Eickhoff, S. B., Thirion, B., Varoquaux, G., & Bzdok, D. (2015). Connectivity-based parcellation: Critique and implications. *Human Brain Mapping*, *36*(12), 4771–4792. https://doi.org/10.1002/HBM.22933

Embleton, K. V., Haroon, H. A., Morris, D. M., Lambon Ralph, M. A., & Parker, G. J. M. (2010). Distortion correction for diffusion-weighted MRI tractography and fMRI in the temporal lobes. *Human Brain Mapping*, *31*(10), 1570–1587. https://doi.org/10.1002/hbm.20959

End, A., & Gamer, M. (2017). Preferential processing of social features and their interplay with physical saliency in complex naturalistic scenes. *Frontiers in Psychology*, *8*(MAR), 418. https://doi.org/10.3389/FPSYG.2017.00418/BIBTEX

REFERENCES

End, A., & Gamer, M. (2019). Task instructions can accelerate the early preference for social features in naturalistic scenes. *Royal Society Open Science*, *6*(3). https://doi.org/10.1098/RSOS.180596

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: Automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience*, *19*(9), 1508–1519. https://doi.org/10.1162/jocn.2007.19.9.1508

Eres, R., Louis, W. R., & Molenberghs, P. (2018). Common and distinct neural networks involved in fMRI studies investigating morality: An ALE meta-analysis. *Social Neuroscience*, *13*(4), 384–398. https://doi.org/10.1080/17470919.2017.1357657

Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Grodd, W., & Wildgruber, D. (2006). Cerebral pathways in processing of affective prosody: A dynamic causal modeling study. *NeuroImage*, *30*(2), 580–587. https://doi.org/10.1016/j.neuroimage.2005.09.059

Evans, J. St. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology*, *59*(1), 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Evans, J. St. B. T. (2011). Dual-process theories of reasoning: Contemporary issues and developmental applications. *Developmental Review*, *31*(2), 86–102. https://doi.org/10.1016/j.dr.2011.07.007

Fan, Y., Duncan, N. W., de Greck, M., & Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience and Biobehavioral Reviews*, *35*(3), 903–911. https://doi.org/10.1016/j.neubiorev.2010.10.009

Faust, M. E., Ferraro, F. R., Balota, D. A., & Spieler, D. H. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin*, *125*(6), 777–799. https://doi.org/10.1037/0033-2909.125.6.777

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(41), 16616–16621. https://doi.org/10.1073/pnas.1315235110

Ferguson, M. J., Mann, T. C., & Wojnowicz, M. T. (2014). Rethinking duality: Criticisms and ways forward. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 578–594). The Guilford Press.

REFERENCES

Fidler, K., & Hütter, M. (2014). The limits of automaticity. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 497–513). The Guilford Press.

Fiedler, M. (1973). Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, *23*(2), 298–305.

Filmer, H. L., Fox, A., & Dux, P. E. (2019). Causal evidence of right temporal parietal junction involvement in implicit Theory of Mind processing. *NeuroImage*, *196*, 329–336. https://doi.org/10.1016/j.neuroimage.2019.04.032

Fingerhut, J., & Prinz, J. J. (2018). Grounding evaluative concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752), 20170142. https://doi.org/10.1098/RSTB.2017.0142

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In *The handbook of social psychology, Vols. 1-2, 4th ed* (pp. 357–411). McGraw-Hill.

Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, *47*, S102. https://doi.org/10.1016/S1053-8119(09)70884-5

Forbes, C. E., & Grafman, J. (2013). Social neuroscience: The second phase. *Frontiers in Human Neuroscience*, *7*(JAN), 20. https://doi.org/10.3389/fnhum.2013.00020

Freeman, J. B., & Johnson, K. L. (2016). More Than Meets the Eye: Split-Second Social Perception. *Trends in Cognitive Sciences*, *20*(5), 362–374. https://doi.org/10.1016/j.tics.2016.03.003

Friederici, A. D. (2011). The Brain Basis of Language Processing: From Structure to Function. *Physiological Reviews*, *91*(4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011

Frith, C. D. (2007). The social brain? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 671–678. https://doi.org/10.1098/RSTB.2006.2003

Frith, C. D., & Frith, U. (2012). Mechanisms of Social Cognition. *Annu. Rev. Psychol*, *63*, 287–313. https://doi.org/10.1146/annurev-psych-120710-100449

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 459–473. https://doi.org/10.1098/rstb.2002.1218

Fujita, K., Trope, Y., Cunningham, W. A., & Liberman, N. (2014). What is control? A conceptual analysis. In J. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 50–68). Guilford Press.

REFERENCES

Fuster, J. (2001). The Prefrontal Cortex—An Update: Time Is of the Essence. *Neuron*, *30*, 319–333.

Gainotti, G. (2015). Is the difference between right and left ATLs due to the distinction between general and social cognition or between verbal and non-verbal representations? *Neuroscience & Biobehavioral Reviews*, *51*, 296–312. https://doi.org/10.1016/J.NEUBIOREV.2015.02.004

Gallese, V., & Lakoff, G. (2007). The Brain's concepts: The role of the Sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3–4), 455–479. https://doi.org/10.1080/02643290442000310

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593–609. https://doi.org/10.1093/brain/119.2.593

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*(9), 396–403. https://doi.org/10.1016/j.tics.2004.07.002

Gao, Z., Zheng, L., Chiou, R., Gouws, A., Krieger-Redwood, K., Wang, X., Varga, D., Lambon Ralph, M. A., Smallwood, J., & Jefferies, E. (2021). Distinct and common neural coding of semantic and non-semantic control demands. *NeuroImage*, *236*, 118230. https://doi.org/10.1016/j.neuroimage.2021.118230

Gazzaniga, M. S. (2014). *Handbook of Cognitive Neuroscience*. Springer.

Gefen, T., Wieneke, C., Martersteck, A., Whitney, K., Weintraub, S., Mesulam, M. M., & Rogalski, E. (2013). Naming vs knowing faces in primary progressive aphasia: A tale of 2 hemispheres. *Neurology*, *81*(7), 658–664. https://doi.org/10.1212/WNL.0b013e3182a08f83

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion*, *12*, 314–325. https://doi.org/10.1037/a0026007

Geng, J. J., & Vossel, S. (2013). Re-evaluating the role of TPJ in attentional control: Contextual updating? *Neuroscience & Biobehavioral Reviews*, *37*(10, Part 2), 2608–2620. https://doi.org/10.1016/j.neubiorev.2013.08.010

Genon, S., Reid, A., Langner, R., Amunts, K., & Eickhoff, S. B. (2018). How to Characterize the Function of a Brain Region. *Trends in Cognitive Sciences*, *22*(4), 350–364. https://doi.org/10.1016/j.tics.2018.01.010

Gerrans, P., & Stone, V. E. (2008). Generous or Parsimonious Cognitive Architecture? Cognitive Neuroscience and Theory of Mind. *The British Journal for the Philosophy of Science*, *59*(2), 121–141. https://doi.org/10.1093/bjps/axm038

REFERENCES

Gilbert, S. J., & Burgess, P. W. (2008). Executive function. *Current Biology*, *18*(3), R110–R114. https://doi.org/10.1016/j.cub.2007.12.014

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

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Van Essen, D. C., & Jenkinson, M. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, *80*, 105–124. https://doi.org/10.1016/J.NEUROIMAGE.2013.04.127

Glenberg, A. M. (2015). Few Believe the World Is Flat: How Embodiment Is Changing the Scientific Understanding of Cognition. *Canadian Journal of Experimental Psychology*, *69*(2), 165–171. https://doi.org/10.1037/CEP0000056

Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex*, *48*(7), 905–922. https://doi.org/10.1016/J.CORTEX.2011.04.010

Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., & Haxby, J. V. (2007). Two takes on the social brain: A comparison of theory of mind tasks. *Journal of Cognitive Neuroscience*, *19*(11), 1803–1814. https://doi.org/10.1162/jocn.2007.19.11.1803

Goghari, V. M., & MacDonald, A. W. (2009). The neural basis of cognitive control: Response selection and inhibition. *Brain and Cognition*, *71*(2), 72–83. https://doi.org/10.1016/j.bandc.2009.04.004

Gold, B. T., Balota, D. A., Jones, S. J., Powell, D. K., Smith, C. D., & Andersen, A. H. (2006). Dissociation of Automatic and Strategic Lexical-Semantics: Functional Magnetic Resonance Imaging Evidence for Differing Roles of Multiple Frontotemporal Regions. *Journal of Neuroscience*, *26*(24), 6523–6532. https://doi.org/10.1523/JNEUROSCI.0808-06.2006

Gonzalez Alam, T., Murphy, C., Smallwood, J., & Jefferies, E. (2018). Meaningful inhibition: Exploring the role of meaning and modality in response inhibition. *NeuroImage*, *181*, 108–119. https://doi.org/10.1016/j.neuroimage.2018.06.074

Gray, K. (2017). How to Map Theory: Reliable Methods Are Fruitless Without Rigorous Theory. *Perspectives on Psychological Science*, *12*(5), 731–741. https://doi.org/10.1177/1745691617691949

REFERENCES

Greven, I. M., & Ramsey, R. (2017). Person perception involves functional integration between the extrastriate body area and temporal pole. *Neuropsychologia*, *96*, 52–60. https://doi.org/10.1016/j.neuropsychologia.2017.01.003

Greven, I. M., Downing, P. E., & Ramsey, R. (2016). Linking person perception and person knowledge in the human brain. *Social Cognitive and Affective Neuroscience*, *11*(4), 641–651. https://doi.org/10.1093/scan/nsv148

Greven, I. M., Downing, P. E., & Ramsey, R. (2019). Neural networks supporting social evaluation of bodies based on body shape. *Social Neuroscience*, *14*(3), 328–344. https://doi.org/10.1080/17470919.2018.1448888

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain Areas Involved in Perception of Biological Motion. *Journal of Cognitive Neuroscience*, *12*(5), 711–720. https://doi.org/10.1162/089892900562417

Haak, K. V., Marquand, A. F., & Beckmann, C. F. (2018). Connectopic mapping with resting-state fMRI. *NeuroImage*, *170*, 83–94. https://doi.org/10.1016/J.NEUROIMAGE.2017.06.075

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423. https://doi.org/10.1016/J.TICS.2005.07.004

Hall, J., Philip, R. C. M., Marwick, K., Whalley, H. C., Romaniuk, L., McIntosh, A. M., Santos, I., Sprengelmeyer, R., Johnstone, E. C., Stanfield, A. C., Young, A. W., & Lawrie, S. M. (2012). Social Cognition, the Male Brain and the Autism Spectrum. *PLoS ONE*, *7*(12), e49033. https://doi.org/10.1371/journal.pone.0049033

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences*, *105*(18), 6741–6746. https://doi.org/10.1073/pnas.0711099105

Happé, F., Cook, J. L., & Bird, G. (2017). The Structure of Social Cognition: In(ter)dependence of Sociocognitive Processes. *Annual Review of Psychology*, *68*(1), 243–267. https://doi.org/10.1146/annurev-psych-010416-044046

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42*(1–3), 335–346. https://doi.org/10.1016/0167-2789(90)90087-6

Harpaintner, M., Trumpp, N. M., & Kiefer, M. (2018). The Semantic Content of Abstract Concepts: A Property Listing Study of 296 Abstract Words. *Frontiers in Psychology*, *9*, 1748. https://doi.org/10.3389/fpsyg.2018.01748

REFERENCES

Hartwigsen, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., & Siebner, H. R. (2010). Phonological decisions require both the left and right supramarginal gyri. *Proceedings of the National Academy of Sciences*, *107*(38), 16494–16499. https://doi.org/10.1073/pnas.1008121107

Hartwigsen, G., Neef, N. E., Camilleri, J. A., Margulies, D. S., & Eickhoff, S. B. (2019). Functional Segregation of the Right Inferior Frontal Gyrus: Evidence From Coactivation-Based Parcellation. *Cerebral Cortex*, *29*(4), 1532–1546. https://doi.org/10.1093/CERCOR/BHY049

Hartwright, C. E., Apperly, I. A., & Hansen, P. C. (2015). The special case of self-perspective inhibition in mental, but not non-mental, representation. *Neuropsychologia*, *67*, 183–192. https://doi.org/10.1016/j.neuropsychologia.2014.12.015

Hassabis, D., Spreng, R. N., Rusu, A. A., Robbins, C. A., Mar, R. A., & Schacter, D. L. (2014). Imagine All the People: How the Brain Creates and Uses Personality Models to Predict Behavior. *Cerebral Cortex*, *24*(8), 1979–1987. https://doi.org/10.1093/cercor/bht042

Haueis, P. (2012). The fuzzy brain. Vagueness and mapping connectivity of the human cerebral cortex. *Frontiers in Neuroanatomy*, *6*. https://doi.org/10.3389/fnana.2012.00037

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

Healey, M. L., & Grossman, M. (2018). Cognitive and Affective Perspective-Taking: Evidence for Shared and Dissociable Anatomical Substrates. *Frontiers in Neurology*, *9*, 491. https://doi.org/10.3389/fneur.2018.00491

Hehman, E., Ingbretsen, Z. A., & Freeman, J. B. (2014). The neural basis of stereotypic impact on multiple social categorization. *NeuroImage*, *101*, 704–711. https://doi.org/10.1016/j.neuroimage.2014.07.056

Henson, R. (2006). Forward inference using functional neuroimaging: Dissociations versus associations. *Trends in Cognitive Sciences*, *10*(2), 64–69. https://doi.org/10.1016/j.tics.2005.12.005

Heyes, C. (2014). Submentalizing: I Am Not Really Reading Your Mind. *Perspectives on Psychological Science*, *9*(2), 131–143. https://doi.org/10.1177/1745691613518076

Heyes, C., & Catmur, C. (2022). What Happened to Mirror Neurons? *Perspectives on Psychological Science*, *17*(1), 153–168. https://doi.org/10.1177/1745691621990638

REFERENCES

Hickok, G. (2009). The functional neuroanatomy of language. *Physics of Life Reviews*, *6*(3), 121–143. https://doi.org/10.1016/j.plrev.2009.06.001

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402. https://doi.org/10.1038/nrn2113

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(6), 1331. https://doi.org/10.1037/0096-1523.22.6.1331

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and Synonymy Effects in Lexical Decision, Naming, and Semantic Categorization Tasks: Interactions between Orthography, Phonology, and Semantics. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*(4), 686–713. https://doi.org/10.1037/0278-7393.28.4.686

Hodges, J. R., & Patterson, K. (2007). Semantic dementia: A unique clinicopathological syndrome. *Lancet Neurology*, *6*(11), 1004–1014. https://doi.org/10.1016/S1474-4422(07)70266-1

Hodgson, C., & Lambon Ralph, M. A. (2008). Mimicking aphasic semantic errors in normal speech production: Evidence from a novel experimental paradigm. *Brain and Language*, *104*(1), 89–101. https://doi.org/10.1016/J.BANDL.2007.03.007

Hodgson, V. J., Lambon Ralph, M. A., & Jackson, R. L. (2021). Multiple dimensions underlying the functional organization of the language network. *NeuroImage*, *241*, 118444. https://doi.org/10.1016/J.NEUROIMAGE.2021.118444

Hoffman, P., & Tamm, A. (2020). Barking up the right tree: Univariate and multivariate fMRI analyses of homonym comprehension. *NeuroImage*, *219*, 117050. https://doi.org/10.1016/j.neuroimage.2020.117050

Hoffman, P., Binney, R. J., & Lambon Ralph, M. A. (2015). Differing contributions of inferior prefrontal and anterior temporal cortex to concrete and abstract conceptual knowledge. *Cortex*, *63*, 250–266. https://doi.org/10.1016/j.cortex.2014.09.001

Hoffman, P., Jefferies, E., & Lambon Ralph, M. A. (2010). Ventrolateral prefrontal cortex plays an executive regulation role in comprehension of abstract words: Convergent neuropsychological and repetitive TMS evidence. *Journal of Neuroscience*, *30*(46), 15450–15456. https://doi.org/10.1523/JNEUROSCI.3783-10.2010

REFERENCES

Hoffman, P., Lambon Ralph, M. A., & Rogers, T. T. (2013). Semantic diversity: A measure of
semantic ambiguity based on variability in the contextual usage of words. *Behavior
Research Methods*, *45*(3), 718–730. https://doi.org/10.3758/s13428-012-0278-x

Hoffman, P., McClelland, J. L., & Lambon Ralph, M. A. (2018). Concepts, control, and
context: A connectionist account of normal and disordered semantic cognition.
Psychological Review, *125*(3), 293–328. https://doi.org/10.1037/rev0000094

Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and
Social Isolation as Risk Factors for Mortality: A Meta-Analytic Review. *Perspectives
on Psychological Science*, *10*(2), 227–237. https://doi.org/10.1177/1745691614568352

Huang, Y., Su, L., & Ma, Q. (2020). The Stroop effect: An activation likelihood estimation
meta-analysis in healthy young adults. *Neuroscience Letters*, *716*, 134683.
https://doi.org/10.1016/j.neulet.2019.134683

Huey, E. D., Krueger, F., & Grafman, J. (2006). Representations in the Human Prefrontal
Cortex. *Current Directions in Psychological Science*, *15*(4), 167–171.
https://doi.org/10.1111/j.1467-8721.2006.00429.x

Hugdahl, K., Raichle, M. E., Mitra, A., & Specht, K. (2015). On the existence of a generalized
non-specific task-dependent network. *Frontiers in Human Neuroscience*, *9*, 1–15.
https://doi.org/10.3389/fnhum.2015.00430

Humphreys, G. F., & Lambon Ralph, M. A. (2015). Fusion and fission of cognitive functions
in the human parietal cortex. *Cerebral Cortex*, *25*(10), 3547–3560.
https://doi.org/10.1093/cercor/bhu198

Humphreys, G. F., Hoffman, P., Visser, M., Binney, R. J., & Lambon Ralph, M. A. (2015).
Establishing task- and modality-dependent dissociations between the semantic and
default mode networks. *Proceedings of the National Academy of Sciences of the United
States of America*, *112*(25), 7857–7862. https://doi.org/10.1073/pnas.1422760112

Humphreys, G.F., & Lambon Ralph, M.A. (2017). Mapping domain-selective and
counterpointed domain-general higher cognitive functions in the lateral parietal cortex:
Evidence from fMRI comparisons of difficulty-varying semantic versus visuo-spatial
tasks, and functional connectivity analyses. *Cerebral Cortex*, *27*(8), 1499–4212.
https://doi.org/10.1093/cercor/bhx107

Humphreys, G., Lambon Ralph, M., & Simons, J. (2021). A Unifying Account of Angular
Gyrus Contributions to Episodic and Semantic Cognition. *Trends in Neurosciences,
44(6), 452-463* https://doi.org/10.1016/j.tins.2021.01.006

REFERENCES

Huntenburg, J. M., Bazin, P. L., & Margulies, D. S. (2018). Large-Scale Gradients in Human Cortical Organization. In *Trends in Cognitive Sciences* (Vol. 22, Issue 1, pp. 21–31). Elsevier Ltd. https://doi.org/10.1016/j.tics.2017.11.002

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453–458. https://doi.org/10.1038/nature17637

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain. *Neuron*, *76*(6), 1210–1224. https://doi.org/10.1016/J.NEURON.2012.10.014

Iacoboni, M. (2008). *Mirroring people: The new science of how we connect with others*. Farrar, Straus and Giroux.

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*, *7*(12), Article 12. https://doi.org/10.1038/nrn2024

Iannetti, G. D., Salomons, T. V., Moayedi, M., Mouraux, A., & Davis, K. D. (2013). Beyond metaphor: Contrasting mechanisms of social and physical pain. *Trends in Cognitive Sciences*, *17*(8), 371–378. https://doi.org/10.1016/j.tics.2013.06.002

Igelström, K. M., & Graziano, M. S. A. (2017). The inferior parietal lobule and temporoparietal junction: A network perspective. *Neuropsychologia*, *105*, 70–83. https://doi.org/10.1016/j.neuropsychologia.2017.01.001

Igelström, K. M., Webb, T. W., Kelly, Y. T., & Graziano, M. S. A. (2016). Topographical Organization of Attentional, Social, and Memory Processes in the Human Temporoparietal Cortex. *ENeuro*, *3*(2), ENEURO.0060-16.2016. https://doi.org/10.1523/ENEURO.0060-16.2016

Irish, M., Hodges, J. R., & Piguet, O. (2014). Right anterior temporal lobe dysfunction underlies theory of mind impairments in semantic dementia. *Brain*, *137*(4), 1241–1253. https://doi.org/10.1093/brain/awu003

Isik, L., Koldewyn, K., Beeler, D., & Kanwisher, N. (2017). Perceiving social interactions in the posterior superior temporal sulcus. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(43), E9145–E9152. https://doi.org/10.1073/PNAS.1714471114/-/DCSUPPLEMENTAL

Jackson, R. L. (2021). The neural correlates of semantic control revisited. *NeuroImage*, *224*, 117444. https://doi.org/10.1016/j.neuroimage.2020.117444

REFERENCES

Jackson, R. L., Bajada, C. J., Lambon Ralph, M. A., & Cloutman, L. L. (2020). The Graded
Change in Connectivity across the Ventromedial Prefrontal Cortex Reveals Distinct
Subregions. *Cerebral Cortex*, *30*(1), 165–180. https://doi.org/10.1093/cercor/bhz079

Jackson, R. L., Bajada, C. J., Rice, G. E., Cloutman, L. L., & Lambon Ralph, M. A. (2018). An
emergent functional parcellation of the temporal cortex. *NeuroImage*, *170*, 385–399.
https://doi.org/10.1016/J.NEUROIMAGE.2017.04.024

Jackson, R. L., Cloutman, L. L., & Lambon Ralph, M. A. (2019). Exploring distinct default
mode and semantic networks using a systematic ICA approach. *Cortex*, *113*, 279–297.
https://doi.org/10.1016/j.cortex.2018.12.019

Jackson, R. L., Hoffman, P., Pobric, G., & Lambon Ralph, M. A. (2016). The Semantic
Network at Work and Rest: Differential Connectivity of Anterior Temporal Lobe
Subregions. *The Journal of Neuroscience*, *36*(5), 1490–1501.
https://doi.org/10.1523/JNEUROSCI.2999-15.2016

Jackson, R. L., Humphreys, G. F., Rice, G. E., Binney, R. J., & Lambon Ralph, M. A. (2021a).
The Coherent Default Mode Network is not involved in Episodic Recall or Social
Cognition. *BioRxiv*, 2021.01.08.425921. https://doi.org/10.1101/2021.01.08.425921

Jackson, R. L., Rogers, T. T., & Lambon Ralph, M. A. (2021b). Reverse-engineering the
cortical architecture for controlled semantic cognition. *Nature Human Behaviour*, 1–13.
https://doi.org/10.1038/s41562-020-01034-z

Jakobsen, E., Böttger, J., Bellec, P., Geyer, S., Rübsamen, R., Petrides, M., & Margulies, D. S.
(2016). Subdivision of Broca's region based on individual-level functional
connectivity. *European Journal of Neuroscience*, *43*(4), 561–571.
https://doi.org/10.1111/ejn.13140

Jakobsen, E., Liem, F., Klados, M. A., Bayrak, Ş., Petrides, M., & Margulies, D. S. (2018).
Automated individual-level parcellation of Broca's region based on functional
connectivity. *NeuroImage*, *170*, 41–53.
https://doi.org/10.1016/j.neuroimage.2016.09.069

January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of stroop and
syntactic ambiguity resolution in Broca's area: Implications for the neural basis of
sentence processing. *Journal of Cognitive Neuroscience*, *21*(12), 2434–2444.
https://doi.org/10.1162/jocn.2008.21179

Jefferies, E. (2013). The neural basis of semantic cognition: Converging evidence from
neuropsychology, neuroimaging and TMS. *Cortex*, *49*(3), 611–625.
https://doi.org/10.1016/j.cortex.2012.10.008

REFERENCES

Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: A case-series comparison. *Brain*, *129*(8), 2132–2147. https://doi.org/10.1093/brain/awl153

Jefferies, E., Baker, S. S., Doran, M., & Lambon Ralph, M. A. (2007). Refractory effects in stroke aphasia: A consequence of poor semantic control. *Neuropsychologia*, *45*(5), 1065–1079. https://doi.org/10.1016/j.neuropsychologia.2006.09.009

Jefferies, E., Patterson, K., & Lambon Ralph, M. A. (2008). Deficits of knowledge versus executive control in semantic cognition: Insights from cued naming. *Neuropsychologia*, *46*(2), 649–658. https://doi.org/10.1016/j.neuropsychologia.2007.09.007

Jeon, H.-A., & Friederici, A. D. (2015). Degree of automaticity and the prefrontal cortex. *Trends in Cognitive Sciences*, *19*(5), 244–250. https://doi.org/10.1016/j.tics.2015.03.003

Jezzard, P., & Clare, S. (1999). Sources of distortion in functional MRI data. *Human Brain Mapping*, *8*(2–3), 80–85. https://doi.org/10.1002/(SICI)1097-0193(1999)8:2/3<80::AID-HBM2>3.0.CO;2-C

Johansen-Berg, H., Behrens, T. E. J., Robson, M. D., Drobnjak, I., Rushworth, M. F. S., Brady, J. M., Smith, S. M., Higham, D. J., & Matthews, P. M. (2004). Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(36), 13335–13340. https://doi.org/10.1073/PNAS.0403743101/SUPPL_FILE/03743FIG8.PDF

Juhasz, B. J., & Yap, M. J. (2012). Sensory experience ratings for over 5,000 mono- and disyllabic words. *Behavior Research Methods 2012 45:1*, *45*(1), 160–168. https://doi.org/10.3758/S13428-012-0242-9

Jung, J. Y., Cloutman, L. L., Binney, R. J., & Lambon Ralph, M. A. (2017). The structural connectivity of higher order association cortices reflects human functional brain networks. *Cortex*, *97*, 221–239. https://doi.org/10.1016/j.cortex.2016.08.011

Jung, J., & Lambon Ralph, M. A. (2022). Distinct but cooperating brain networks supporting semantic cognition. *Cerebral Cortex*, bhac190. https://doi.org/10.1093/cercor/bhac190

Kamminga, J., Kumfor, F., Burrell, J. R., Piguet, O., Hodges, J. R., & Irish, M. (2015). Differentiating between right-lateralised semantic dementia and behavioural-variant frontotemporal dementia: An examination of clinical characteristics and emotion processing. *Journal of Neurology, Neurosurgery &Amp; Psychiatry*, *86*(10), 1082 LP – 1088. https://doi.org/10.1136/jnnp-2014-309120

REFERENCES

Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(25), 11163–11170. https://doi.org/10.1073/PNAS.1005062107/-/DCSUPPLEMENTAL

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109–2128. https://doi.org/10.1098/RSTB.2006.1934

Keller, S. S., Crow, T., Foundas, A., Amunts, K., & Roberts, N. (2009). Broca's area: Nomenclature, anatomy, typology and asymmetry. *Brain and Language*, *109*(1), 29–48. https://doi.org/10.1016/J.BANDL.2008.11.005

Kelly, C., Uddin, L. Q., Shehzad, Z., Margulies, D. S., Castellanos, F. X., Milham, M. P., & Petrides, M. (2010). Broca's region: Linking human brain functional connectivity data and non-human primate tracing anatomy studies. *European Journal of Neuroscience*, *32*(3), 383–398. https://doi.org/10.1111/j.1460-9568.2010.07279.x

Kiebel, S. J., Daunizeau, J., & Friston, K. J. (2008). A Hierarchy of Time-Scales and the Brain. *PLOS Computational Biology*, *4*(11), e1000209. https://doi.org/10.1371/journal.pcbi.1000209

Kiefer, M., & Harpaintner, M. (2020). Varieties of abstract concepts and their grounding in perception or action. *Open Psychology*, *2*(1), 119–137. https://doi.org/10.1515/PSYCH-2020-0104

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, *48*(7), 805–825. https://doi.org/10.1016/j.cortex.2011.04.006

Kilner, J. M. (2011). More than one pathway to action understanding. *Trends in Cognitive Sciences*, *15*(8), 352–357. https://doi.org/10.1016/j.tics.2011.06.005

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of Mirror Neurons in Human Inferior Frontal Gyrus. *Journal of Neuroscience*, *29*(32), 10153–10159. https://doi.org/10.1523/JNEUROSCI.2668-09.2009

Kittel, A. F. D., Olderbak, S., & Wilhelm, O. (2022). Sty in the Mind's Eye: A Meta-Analytic Investigation of the Nomological Network and Internal Consistency of the "Reading the Mind in the Eyes" Test. *Assessment*, *29*(5), 872–895. https://doi.org/10.1177/1073191121996469

Klein, J. C., Behrens, T. E. J., Robson, M. D., Mackay, C. E., Higham, D. J., & Johansen-Berg, H. (2007). Connectivity-based parcellation of human cortex using diffusion MRI:

REFERENCES

Establishing reproducibility, validity and observer independence in BA 44/45 and SMA/pre-SMA. *NeuroImage*, *34*(1), 204–211. https://doi.org/10.1016/j.neuroimage.2006.08.022

Klüver, H., & Bucy, P. C. (1939). Preliminary analysis of functions of the temporal lobes in monkeys. *Archives of Neurology & Psychiatry*, *42*, 979–1000. https://doi.org/10.1001/archneurpsyc.1939.02270240017001

Koechlin, E., & Jubault, T. (2006). Broca's Area and the Hierarchical Organization of Human Behavior. *Neuron*, *50*(6), 963–974. https://doi.org/10.1016/j.neuron.2006.05.017

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

Korman, J., & Malle, B. F. (2016). Grasping for Traits or Reasons? How People Grapple With Puzzling Social Behaviors. *Personality and Social Psychology Bulletin*, *42*(11), 1451–1465. https://doi.org/10.1177/0146167216663704

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The Representation of Abstract Words: Why Emotion Matters. *Journal of Experimental Psychology: General*, *140*(1), 14–34. https://doi.org/10.1037/a0021446

Krall, S. C., Volz, L. J., Oberwelland, E., Grefkes, C., Fink, G. R., & Konrad, K. (2016). The right temporoparietal junction in attention and social interaction: A transcranial magnetic stimulation study. *Human Brain Mapping*, *37*(2), 796–807. https://doi.org/10.1002/hbm.23068

Krieger-Redwood, K., Teige, C., Davey, J., Hymers, M., & Jefferies, E. (2015). Conceptual control across modalities: Graded specialisation for pictures and words in inferior frontal and posterior temporal cortex. *Neuropsychologia*, *76*, 92–107. https://doi.org/10.1016/j.neuropsychologia.2015.02.030

Kuhnke, P., Beaupain, M., Arola, J., Kiefer, M., & Hartwigsen, G. (2022). *Meta-analytic evidence for a novel hierarchical model of conceptual processing*. https://doi.org/10.1101/2022.11.05.515278

Kuhnke, P., Kiefer, M., & Hartwigsen, G. (2020). Task-Dependent Recruitment of Modality-Specific and Multimodal Regions during Conceptual Processing. *Cerebral Cortex (New York, N.Y. : 1991)*, *30*(7), 3938–3959. https://doi.org/10.1093/cercor/bhaa010

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

REFERENCES

Laird, A. R., Eickhoff, S. B., Rottschy, C., Bzdok, D., Ray, K. L., & Fox, P. T. (2013). Networks of task co-activations. *NauroImage*, *80*, 505–514. https://doi.org/10.1016/j.neuroimage.2013.04.073

Lambon Ralph, M. A., & Patterson, K. (2008). Generalization and Differentiation in Semantic Memory Insights from Semantic Dementia. *Annals of the New York Academy of Sciences*, *1124*, 61–76. https://doi.org/10.1196/annals.1440.006

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42–55. https://doi.org/10.1038/nrn.2016.150

Lambon Ralph, M. A., Lowe, C., & Rogers, T. T. (2007). Neural basis of category-specific semantic deficits for living things: Evidence from semantic dementia, HSVE and a neural network model. *Brain*, *130*(4), 1127–1137. https://doi.org/10.1093/brain/awm025

Lambon Ralph, M. A., Sage, K., Jones, R. W., & Mayberry, E. J. (2010). Coherent concepts are computed in the anterior temporal lobes. *Proceedings of the National Academy of Sciences*, *107*(6), 2717–2722. https://doi.org/10.1073/pnas.0907307107

Lamm, C., Bukowski, H., & Silani, G. (2016). From shared to distinct self–other representations in empathy: Evidence from neurotypical function and socio-cognitive disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150083. https://doi.org/10.1098/rstb.2015.0083

Lancaster, J. L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., Mazziotta, J. C., & Fox, P. T. (2007). Bias between MNI and talairach coordinates analyzed using the ICBM-152 brain template. *Human Brain Mapping*, *28*(11), 1194–1205. https://doi.org/10.1002/hbm.20345

Lavoie, M. A., Vistoli, D., Sutliff, S., Jackson, P. L., & Achim, A. M. (2016). Social representations and contextual adjustments as two distinct components of the Theory of Mind brain network: Evidence from the REMICS task. *Cortex*, *81*, 176–191. https://doi.org/10.1016/j.cortex.2016.04.017

Lee, S. M., & McCarthy, G. (2016). Functional Heterogeneity and Convergence in the Right Temporoparietal Junction. *Cerebral Cortex*, *26*(3), 1108–1116. https://doi.org/10.1093/cercor/bhu292

Lega, C., Santandrea, E., Ferrante, O., Serpe, R., Dolci, C., Baldini, E., Cattaneo, L., & Chelazzi, L. (2020). Modulating the influence of recent trial history on attentional

REFERENCES

capture via transcranial magnetic stimulation (TMS) of right TPJ. *Cortex*, *133*, 149–160. https://doi.org/10.1016/j.cortex.2020.09.009

Leshinskaya, A., Contreras, J. M., Caramazza, A., & Mitchell, J. P. (2017). Neural Representations of Belief Concepts: A Representational Similarity Approach to Social Semantics. *Cerebral Cortex*, *27*(1), 344–357. https://doi.org/10.1093/CERCOR/BHW401

Leslie, A. M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, *50*(1), 211–238. https://doi.org/10.1016/0010-0277(94)90029-9

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences*, *8*(12), 528–533. https://doi.org/10.1016/j.tics.2004.10.001

Lewis, M. L., & Frank, M. C. (2016). The length of words reflects their conceptual complexity. *Cognition*, *153*, 182–195. https://doi.org/10.1016/J.COGNITION.2016.04.003

Libero, L. E., Maximo, J. O., Deshpande, H. D., Klinger, L. G., Klinger, M. K., & Kana, R. K. (2014). The role of mirroring and mentalizing networks in mediating action intentions in autism. *Molecular Autism*, *5*(1), 50. https://doi.org/10.1186/2040-2392-5-50

Lieberman, M. D. (2007). Social Cognitive Neuroscience: A Review of Core Processes. *Annual Review of Psychology*, *58*(1), 259–289. https://doi.org/10.1146/annurev.psych.58.110405.085654

Lin, N., Bi, Y., Zhao, Y., Luo, C., & Li, X. (2015). The theory-of-mind network in support of action verb comprehension: Evidence from an fMRI study. *Brain and Language*, *141*, 1–10. https://doi.org/10.1016/J.BANDL.2014.11.004

Lin, N., Wang, X., Xu, Y., Wang, X., Hua, H., Zhao, Y., & Li, X. (2018). Fine Subdivisions of the Semantic Network Supporting Social and Sensory–Motor Semantic Processing. *Cerebral Cortex*, *28*(8), 2699–2710. https://doi.org/10.1093/CERCOR/BHX148

Lin, N., Xu, Y., Wang, X., Yang, H., Du, M., Hua, H., & Li, X. (2019). Coin, telephone, and handcuffs: Neural correlates of social knowledge of inanimate objects. *Neuropsychologia*, *133*, UNSP 107187. https://doi.org/10.1016/j.neuropsychologia.2019.107187

Lin, N., Xu, Y., Yang, H., Zhang, G., Zhang, M., Wang, S., Hua, H., & Li, X. (2020). Dissociating the neural correlates of the sociality and plausibility effects in simple conceptual combination. *Brain Structure and Function*, *225*(3), 995–1008. https://doi.org/10.1007/S00429-020-02052-3

REFERENCES

Lin, N., Yang, X., Li, J., Wang, S., Hua, H., Ma, Y., & Li, X. (2018). Neural correlates of three
cognitive processes involved in theory of mind and discourse comprehension.
Cognitive, Affective and Behavioral Neuroscience, *18*(2), 273–283.
https://doi.org/10.3758/s13415-018-0568-6

Lindquist, K. A., Gendron, M., Barrett, L. F., & Dickerson, B. C. (2014). Emotion perception,
but not affect perception, is impaired with semantic memory loss. *Emotion*, *14*(2), 375–
387. https://doi.org/10.1037/a0035293

Luzzi, S., Snowden, J. S., Neary, D., Coccia, M., Provinciali, L., & Lambon Ralph, M. A.
(2007). Distinct patterns of olfactory impairment in Alzheimer's disease, semantic
dementia, frontotemporal dementia, and corticobasal degeneration. *Neuropsychologia*,
45(8), 1823–1831. https://doi.org/10.1016/j.neuropsychologia.2006.12.008

Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster
Sensorimotor Norms: Multidimensional measures of perceptual and action strength for
40,000 English words. *Behavior Research Methods*, *52*, 1271–1291.
https://doi.org/10.3758/s13428-019-01316-z

Ma, N., Vandekerckhove, M., Baetens, K., Overwalle, F. Van, Seurinck, R., & Fias, W.
(2012). Inconsistencies in spontaneous and intentional trait inferences. *Social Cognitive
and Affective Neuroscience*, *7*(8), 937–950. https://doi.org/10.1093/scan/nsr064

Majerus, S., Péters, F., Bouffier, M., Cowan, N., & Phillips, C. (2018). *The Dorsal Attention
Network Reflects Both Encoding Load and Top–down Control during Working
Memory*. *30*(2), 144–159. https://doi.org/10.1162/jocn_a_01195

Malle, B., & Holbrook, J. (2012). Is There a Hierarchy of Social Inferences? The likelihood
and speed of inferring intentionality, mind, and personality. *Article in Journal of
Personality and Social Psychology*, *102*(4), 661–684. https://doi.org/10.1037/a0026790

Malt, B. C., & Majid, A. (2013). How thought is mapped into words. *Wiley Interdisciplinary
Reviews: Cognitive Science*, *4*(6), 583–597. https://doi.org/10.1002/WCS.1251

Mandera, P., Keuleers, E., & Brysbaert, M. (2020). Recognition times for 62 thousand English
words: Data from the English Crowdsourcing Project. *Behavior Research Methods*,
52(2), 741–760. https://doi.org/10.3758/s13428-019-01272-8

Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G.,
Bezgin, G., Eickhoff, S. B., Castellanos, F. X., Petrides, M., Jefferies, E., &
Smallwood, J. (2016). Situating the default-mode network along a principal gradient of
macroscale cortical organization. *Proceedings of the National Academy of Sciences of*

the United States of America, *113*(44), 12574–12579.
https://doi.org/10.1073/pnas.1608282113

Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., & Halgren, E. (2003).
Spatiotemporal Dynamics of Modality-Specific and Supramodal Word Processing.
Neuron, *38*(3), 487–497. https://doi.org/10.1016/S0896-6273(03)00197-1

Mars, R. B., Neubert, F.-X., Noonan, M. P., Sallet, J., Toni, I., & Rushworth, M. F. S. (2012).
On the relationship between the "default mode network" and the "social brain".
Frontiers in Human Neuroscience, *6*, 1–9. https://doi.org/10.3389/fnhum.2012.00189

Mars, R. B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., & Rushworth, M. F. S. (2012).
Connectivity-Based Subdivisions of the Human Right "Temporoparietal Junction
Area": Evidence for Different Areas Participating in Different Cortical Networks.
Cerebral Cortex, *22*(8), 1894–1903. https://doi.org/10.1093/cercor/bhr268

Mason, R. A., Williams, D. L., Kana, R. K., Minshew, N., & Just, M. A. (2008). Theory of
Mind disruption and recruitment of the right hemisphere during narrative
comprehension in autism. *Neuropsychologia*, *46*(1), 269–280.
https://doi.org/10.1016/j.neuropsychologia.2007.07.018

Maurage, P., Grynberg, D., Noël, X., Joassin, F., Hanak, C., Verbanck, P., Luminet, O., de
Timary, P., Campanella, S., & Philippot, P. (2011). The "Reading the Mind in the
Eyes" test as a new way to explore complex emotions decoding in alcohol dependence.
Psychiatry Research, *190*(2), 375–378. https://doi.org/10.1016/j.psychres.2011.06.015

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houdé, O., Crivello, F., Joliot, M.,
Petit, L., & Tzourio-Mazoyer, N. (2001). Cortical networks for working memory and
executive functions sustain the conscious resting state in man. *Brain Research Bulletin*,
54(3), 287–298. https://doi.org/10.1016/S0361-9230(00)00437-8

Mazzuca, C., Majid, A., Lugli, L., Nicoletti, R., & Borghi, A. M. (2020). Gender is a
multifaceted concept: Evidence that specific life experiences differentially shape the
concept of gender. *Language and Cognition*, *12*(4), 649–678.
https://doi.org/10.1017/langcog.2020.15

McDonald, S. (2013). Impairments in Social Cognition Following Severe Traumatic Brain
Injury. *Journal of the International Neuropsychological Society*, *19*(3), 231–246.
https://doi.org/10.1017/S1355617712001506

Mellem, M. S., Jasmin, K. M., Peng, C., & Martin, A. (2016). Sentence processing in anterior
superior temporal cortex shows a social-emotional bias. *Neuropsychologia*, *89*, 217–
224. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2016.06.019

REFERENCES

Melnikoff, D. E., & Bargh, J. A. (2018). The Mythical Number Two. In *Trends in Cognitive Sciences* (Vol. 22, Issue 4, pp. 280–293). Elsevier Ltd. https://doi.org/10.1016/j.tics.2018.02.001

Mende-Siedlecki, P., & Todorov, A. (2016). Neural dissociations between meaningful and mere inconsistency in impression updating. *Social Cognitive and Affective Neuroscience*, *11*(9), 1489–1500. https://doi.org/10.1093/scan/nsw058

Mende-Siedlecki, P., Baron, S. G., & Todorov, A. (2013). Diagnostic value underlies asymmetric updating of impressions in the morality and ability domains. *Journal of Neuroscience*, *33*(50), 19406–19415. https://doi.org/10.1523/JNEUROSCI.2334-13.2013

Mende-Siedlecki, P., Cai, Y., & Todorov, A. (2013). The neural dynamics of updating person impressions. *Social Cognitive and Affective Neuroscience*, *8*(6), 623–631. https://doi.org/10.1093/scan/nss040

Mende-siedlecki, P., Said, C. P., & Todorov, A. (2013). The social evaluation of faces: A meta-analysis of functional neuroimaging studies. *Social Cognitive and Affective Neuroscience*, *8*(3), 285–299. https://doi.org/10.1093/scan/nsr090

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: A network model of insula function. *Brain Structure and Function*, *214*(5), 655–667. https://doi.org/10.1007/s00429-010-0262-0

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, *48*(7), 788–804. https://doi.org/10.1016/j.cortex.2010.11.002

Meyer, M. L., Spunt, R. P., Berkman, E. T., Taylor, S. E., & Lieberman, M. D. (2012). Evidence for social working memory from a parametric functional MRI study. *PNAS*, *109*(6), 1883–1888. https://doi.org/10.1073/pnas.1121077109

Meyer, M. L., Taylor, S. E., & Lieberman, M. D. (2015). Social working memory and its distinctive link to social cognitive ability: An fMRI study. *Social Cognitive and Affective Neuroscience*, *10*(10), 1338–1347. https://doi.org/10.1093/scan/nsv065

Meyer, M., & Lieberman, M. (2012). Social Working Memory: Neurocognitive Networks and Directions for Future Research. *Frontiers in Psychology*, *3*. https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00571

Michael, J., & D'Ausilio, A. (2015). Domain-specific and domain-general processes in social perception – A complementary approach. *Consciousness and Cognition*, *36*, 434–437. https://doi.org/10.1016/j.concog.2014.12.009

REFERENCES

Mier, D., Sauer, C., Lis, S., Esslinger, C., Wilhelm, J., Gallhofer, B., & Kirsch, P. (2010). Neuronal correlates of affective theory of mind in schizophrenia out-patients: Evidence for a baseline deficit. *Psychological Medicine*, *40*(10), 1607–1617. https://doi.org/10.1017/S0033291709992133

Mitchell, J. P. (2008). Activity in Right Temporo-Parietal Junction is Not Selective for Theory-of-Mind. *Cerebral Cortex*, *18*(2), 262–271. https://doi.org/10.1093/cercor/bhm051

Mitchell, R. L. C. (2013). Further characterisation of the functional neuroanatomy associated with prosodic emotion decoding. *Cortex*, *49*(6), 1722–1732. https://doi.org/10.1016/j.cortex.2012.07.010

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2012). Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies. *Neuroscience & Biobehavioral Reviews*, *36*(1), 341–349. https://doi.org/10.1016/j.neubiorev.2011.07.004

Molenberghs, P., Johnson, H., Henry, J. D., & Mattingley, J. B. (2016). Understanding the minds of others: A neuroimaging meta-analysis. *Neuroscience & Biobehavioral Reviews*, *65*, 276–291. https://doi.org/10.1016/j.neubiorev.2016.03.020

Muhle-Karbe, P. S., Derrfuss, J., Lynn, M. T., Neubert, F. X., Fox, P. T., Brass, M., & Eickhoff, S. B. (2016). Co-Activation-Based Parcellation of the Lateral Prefrontal Cortex Delineates the Inferior Frontal Junction Area. *Cerebral Cortex*, *26*(5), 2225–2241. https://doi.org/10.1093/cercor/bhv073

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-Neuron Responses in Humans during Execution and Observation of Actions. *Current Biology*, *20*(8), 750–756. https://doi.org/10.1016/j.cub.2010.02.045

Müller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., Tench, C. R., Yarkoni, T., Nichols, T. E., Turkeltaub, P. E., Wager, T. D., & Eickhoff, S. B. (2018). Ten simple rules for neuroimaging meta-analysis. *Neuroscience and Biobehavioral Reviews*, *84*, 151–161. https://doi.org/10.1016/j.neubiorev.2017.11.012

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., Simonsohn, U., Wagenmakers, E. J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. In *Nature Human Behaviour* (Vol. 1, Issue 1, pp. 1–9). Nature Publishing Group. https://doi.org/10.1038/s41562-016-0021

Muraki, E. J., Sidhu, D. M., & Pexman, P. M. (2019). Mapping semantic space: Property norms and semantic richness. *Cognitive Processing*. https://doi.org/10.1007/s10339-019-00933-y

REFERENCES

Muraki, E. J., Sidhu, D. M., & Pexman, P. M. (2020). Heterogenous abstract concepts: Is "ponder" different from "dissolve"? *Psychological Research*. https://doi.org/10.1007/s00426-020-01398-x

Murphy, C., Rueschemeyer, S.-A., Watson, D., Karapanagiotidis, T., Smallwood, J., & Jefferies, E. (2017). Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality. *NeuroImage*, *147*, 19–31. https://doi.org/10.1016/j.neuroimage.2016.11.067

Nagel, I. E., Schumacher, E. H., Goebel, R., & D'Esposito, M. (2008). Functional MRI investigation of verbal selection mechanisms in lateral prefrontal cortex. *NeuroImage*, *43*(4), 801–807. https://doi.org/10.1016/j.neuroimage.2008.07.017

Nakae, T., Matsumoto, R., Kunieda, T., Arakawa, Y., Kobayashi, K., Shimotake, A., Yamao, Y., Kikuchi, T., Aso, T., Matsuhashi, M., Yoshida, K., Ikeda, A., Takahashi, R., Lambon Ralph, M. A., & Miyamoto, S. (2020). Connectivity Gradient in the Human Left Inferior Frontal Gyrus: Intraoperative Cortico-Cortical Evoked Potential Study. *Cerebral Cortex*, *30*(8), 4633–4650. https://doi.org/10.1093/cercor/bhaa065

Nee, D. E., & D'Esposito, M. (2016). The hierarchical organization of the lateral prefrontal cortex. *ELife*, *5*, e12112. https://doi.org/10.7554/eLife.12112

Nee, D. E., Wager, T. D., & Jonides, J. (2007). Interference resolution: Insights from a meta-analysis of neuroimaging tasks. *Cognitive, Affective and Behavioral Neuroscience*, *7*(1), 1–17. https://doi.org/10.3758/CABN.7.1.1

Nestor, A., Plaut, D. C., & Behrmann, M. (2011). Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proceedings of the National Academy of Sciences*, *108*(24), 9998–10003. https://doi.org/10.1073/pnas.1102433108

Neubert, F. X., Mars, R. B., Thomas, A. G., Sallet, J., & Rushworth, M. F. S. (2014). Comparison of Human Ventral Frontal Cortex Areas for Cognitive Control and Language with Areas in Monkey Frontal Cortex. *Neuron*, *81*(3), 700–713. https://doi.org/10.1016/j.neuron.2013.11.012

Newcombe, P. I., Campbell, C., Siakaluk, P. D., & Pexman, P. M. (2012). Effects of Emotional and Sensorimotor Knowledge in Semantic Processing of Concrete and Abstract Nouns. *Frontiers in Human Neuroscience*, *6*, 275. https://doi.org/10.3389/fnhum.2012.00275

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage*, *25*(3), 653–660. https://doi.org/10.1016/j.neuroimage.2004.12.005

REFERENCES

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(2), 241–268. https://doi.org/10.3758/s13415-011-0083-5

Nook, E. C., Lindquist, K. A., & Zaki, J. (2015). A new look at emotion perception: Concepts speed and shape facial emotion recognition. *Emotion*, *15*, 569–578. https://doi.org/10.1037/a0039166

Nook, E. C., Stavish, C. M., Sasse, S. F., Lambert, H. K., Mair, P., McLaughlin, K. A., & Somerville, L. H. (2020). Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures. *Emotion*, *20*(5), 773–792. https://doi.org/10.1037/EMO0000609

Noonan, K. A., Jefferies, E., Corbett, F., & Lambon Ralph, M. A. (2010). Elucidating the Nature of Deregulated Semantic Cognition in Semantic Aphasia: Evidence for the Roles of Prefrontal and Temporo-parietal Cortices. *Journal of Cognitive Neuroscience*, *22*(7), 1597–1613. https://doi.org/10.1162/jocn.2009.21289

Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013). Going beyond Inferior Prefrontal Involvement in Semantic Control: Evidence for the Additional Contribution of Dorsal Angular Gyrus and Posterior Middle Temporal Cortex. *Journal of Cognitive Neuroscience*, *25*(11), 1824–1850. https://doi.org/10.1162/jocn_a_00442

Numssen, O., Bzdok, D., & Hartwigsen, G. (2021). Functional specialization within the inferior parietal lobes across cognitive domains. *ELife*, *10*. https://doi.org/10.7554/eLife.63591

Ojemann, J. G., Akbudak, E., Snyder, A. Z., McKinstry, R. C., Raichle, M. E., & Conturo, T. E. (1997). Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. *NeuroImage*, *6*(3), 156–167. https://doi.org/10.1006/nimg.1997.0289

Olderbak, S., Geiger, M., & Wilhelm, O. (2019). A call for revamping socio-emotional ability research in autism. *Behavioral and Brain Sciences*, *42*. https://doi.org/10.1017/S0140525X1800239X

Olson, I. R., Mccoy, D., Klobusicky, E., & Ross, L. A. (2013). Social cognition and the anterior temporal lobes: A review and theoretical framework. *Social Cognitive and Affective Neuroscience*, *8*, 123–133. https://doi.org/10.1093/scan/nss119

REFERENCES

Ortiz-Tudela, J., Martín-Arévalo, E., Chica, A. B., & Lupiáñez, J. (2018). Semantic incongruity attracts attention at a pre-conscious level: Evidence from a TMS study. *Cortex*, *102*, 96–106. https://doi.org/10.1016/j.cortex.2017.08.035

Özdem, C., Brass, M., Van der Cruyssen, L., & Van Overwalle, F. (2017). The overlap between false belief and spatial reorientation in the temporo-parietal junction: The role of input modality and task. *Social Neuroscience*, *12*(2), 207–217. https://doi.org/10.1080/17470919.2016.1143027

Papitto, G., Friederici, A. D., & Zaccarella, E. (2020). The topographical organization of motor processing: An ALE meta-analysis on six action domains and the relevance of Broca's region. *NeuroImage*, *206*, 116321. https://doi.org/10.1016/J.NEUROIMAGE.2019.116321

Passingham, R. E., Stephan, K. E., & Kötter, R. (2002). The anatomical basis of functional localization in the cortex. *Nature Reviews Neuroscience 2002 3:8*, *3*(8), 606–616. https://doi.org/10.1038/nrn893

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976–987. https://doi.org/10.1038/nrn2277

Peelen, M. V., & Caramazza, A. (2012). Conceptual Object Representations in Human Anterior Temporal Cortex. *Journal of Neuroscience*, *32*(45), 15728–15736. https://doi.org/10.1523/JNEUROSCI.1953-12.2012

Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience*, *8*(8), Article 8. https://doi.org/10.1038/nrn2195

Pérez-Edgar, K., & Fox, N. A. (2007). Temperamental contributions to children's performance in an emotion-word processing task: A behavioral and electrophysiological study. *Brain and Cognition*, *65*(1), 22–35. https://doi.org/10.1016/J.BANDC.2006.10.010

Perner, J., & Leekam, S. (2008). The Curious Incident of the Photo that was Accused of Being False: Issues of Domain Specificity in Development, Autism, and Brain Imaging. *Quarterly Journal of Experimental Psychology*, *61*(1), 76–89. https://doi.org/10.1080/17470210701508756

Perry, R. J., Rosen, H. R., Kramer, J. H., Beer, J. S., Levenson, R. L., & Miller, B. L. (2001). Hemispheric Dominance for Emotions, Empathy and Social Behaviour: Evidence from Right and Left Handers with Frontotemporal Dementia. *Neurocase*, *7*(2), 145–160. https://doi.org/10.1093/NEUCAS/7.2.145

REFERENCES

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, *35*, 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525

Petrides, M. (2005). Lateral prefrontal cortex: Architectonic and functional organization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 781–795. https://doi.org/10.1098/rstb.2005.1631

Pexman, P. M. (2008). It's Fascinating Research: The Cognition of Verbal Irony. *Current Directions in Psychological Science*, *17*(4), 286–290. https://doi.org/10.1111/j.1467-8721.2008.00591.x

Pexman, P. M. (2012). Meaning-level influences on visual word recognition. In J. Adelman (Ed.), *Visual Word Recognition: Meaning and context, individuals and development* (pp. 24–43). Psychology Press.

Pexman, P. M. (2020). How Does Meaning Come to Mind? Four Broad Principles of Semantic Processing. *Canadian Journal of Experimental Psychology*, *74*(4), 275–283. https://doi.org/10.1037/CEP0000235

Pexman, P. M., & Yap, M. J. (2018). Individual differences in semantic processing: Insights from the Calgary Semantic Decision Project. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*(7), 1091–1112. https://doi.org/10.1037/XLM0000499

Pexman, P. M., Diveica, V., & Binney, R. J. (2021). Social Semantics: The Organisation and Grounding of Abstract Concepts. *PsyArXiv*. https://doi.org/10.31234/OSF.IO/WRBGP

Pexman, P. M., Heard, A., Lloyd, E., & Yap, M. J. (2017). The Calgary semantic decision project: Concrete/abstract decision data for 10,000 English words. *Behavior Research Methods*, *49*(2), 407–417. https://doi.org/10.3758/s13428-016-0720-6

Pexman, P. M., Lupker, S. J., & Hino, Y. (2002). The impact of feedback semantics in visual word recognition: Number-of-features effects in lexical decision and naming tasks. *Psychonomic Bulletin & Review*, *9*(3), 542–549. https://doi.org/10.3758/BF03196311

Pexman, P. M., Muraki, E., Sidhu, D. M., Siakaluk, P. D., & Yap, M. J. (2019). Quantifying sensorimotor experience: Body–object interaction ratings for more than 9,000 English words. *Behavior Research Methods*, *51*(2), 453–466. https://doi.org/10.3758/s13428-018-1171-z

Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a Third Visual Pathway Specialized for Social Perception. *Trends in Cognitive Sciences*, *25*(2), 100–110. https://doi.org/10.1016/J.TICS.2020.11.006

REFERENCES

Piwnica Worms, K. E., Omar, R., Hailstone, J. C., & Warren, J. D. (2010). Flavour processing in semantic dementia. *Cortex*, *46*(6), 761–768. https://doi.org/10.1016/j.cortex.2009.07.002

Plaut, D. C. (2002). Graded modality-specific specialisation in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology*, *19*(7), 603–639. https://doi.org/10.1080/02643290244000112

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2007). Anterior temporal lobes mediate semantic representation: Mimicking semantic dementia by using rTMS in normal participants. *PNAS*, *104*(50), 20137–20141.

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010a). Amodal semantic representations depend on both anterior temporal lobes: Evidence from repetitive transcranial magnetic stimulation. *Neuropsychologia*, *48*(5), 1336–1342. https://doi.org/10.1016/j.neuropsychologia.2009.12.036

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010b). Category-Specific versus Category-General Semantic Impairment Induced by Transcranial Magnetic Stimulation. *Current Biology*, *20*(10), 964–968. https://doi.org/10.1016/j.cub.2010.03.070

Pobric, G., Lambon Ralph, M. A., & Jefferies, E. (2009). The role of the anterior temporal lobes in the comprehension of concrete and abstract words: RTMS evidence. *Cortex*, *45*(9), 1104–1110. https://doi.org/10.1016/j.cortex.2009.02.006

Pobric, G., Lambon Ralph, M., & Zahn, R. (2016). Hemispheric Specialization within the Superior Anterior Temporal Cortex for Social and Nonsocial Concepts. *Journal of Cognitive Neuroscience*, *28*(3), 351–360. https://doi.org/10.1162/JOCN_A_00902

Poeppel, D., Mangun, G. R., & Gazzaniga, M. S. (2020). *The Cognitive Neurosciences, sixth edition*. MIT Press.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional Specialization for Semantic and Phonological Processing in the Left Inferior Prefrontal Cortex. *NeuroImage*, *10*(1), 15–35. https://doi.org/10.1006/nimg.1999.0441

Powell, J. L., Grossi, D., Corcoran, R., Gobet, F., & García-Fiñana, M. (2017). The neural correlates of theory of mind and their role during empathy and the game of chess: A functional magnetic resonance imaging study. *Neuroscience*, *355*, 149–160. https://doi.org/10.1016/j.neuroscience.2017.04.042

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

REFERENCES

Press, C., Weiskopf, N., & Kilner, J. M. (2012). Dissociable roles of human inferior frontal gyrus during action execution and observation. *NeuroImage*, *60*(3), 1671–1677. https://doi.org/10.1016/j.neuroimage.2012.01.118

Price, C. J. (2012). A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, *62*(2), 816–847. https://doi.org/10.1016/j.neuroimage.2012.04.062

Pulvermüller, F. (2001). Brain reflections of words and their meaning. *Trends in Cognitive Sciences*, *5*(12), 517–524. https://doi.org/10.1016/S1364-6613(00)01803-9

Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. MIT Press.

Quadflieg, S., & Koldewyn, K. (2017). The neuroscience of people watching: How the human brain makes sense of other people's encounters. *Annals of the New York Academy of Sciences*, *1396*(1), 166–182. https://doi.org/10.1111/nyas.13331

Qualtrics. (2020). https://www.qualtrics.com/

Quesque, F., & Brass, M. (2019). The Role of the Temporoparietal Junction in Self-Other Distinction. *Brain Topography*, *32*(6), 943–955. https://doi.org/10.1007/s10548-019-00737-5

Quesque, F., & Rossetti, Y. (2020). What Do Theory-of-Mind Tasks Actually Measure? Theory and Practice. *Perspectives on Psychological Science*, *15*(2), 384–396. https://doi.org/10.1177/1745691619896607

Quillian, M. R. (1969). The teachable language comprehender: A simulation program and theory of language. *Communications of the the Association for Computing Machinery*, *12*(8), 459–476. https://doi.org/10.1145/363196.363214

Quinn, K. A., & Rosenthal, H. E. S. (2012). Categorizing others and the self: How social memory structures guide social perception and behavior. *Learning and Motivation*, *43*(4), 247–258. https://doi.org/10.1016/j.lmot.2012.05.008

Qureshi, A. W., Bretherton, L., Marsh, B., & Monk, R. L. (2020). Stimulation of the dorsolateral prefrontal cortex impacts conflict resolution in Level-1 visual perspective taking. *Cognitive, Affective, & Behavioral Neuroscience*, *20*(3), 565–574. https://doi.org/10.3758/s13415-020-00786-5

Rämä, P., & Courtney, S. M. (2005). Functional topography of working memory for face or voice identity. *NeuroImage*, *24*(1), 224–234. https://doi.org/10.1016/j.neuroimage.2004.08.024

REFERENCES

Rämä, P., Martinkauppi, S., Linnankoski, I., Koivisto, J., Aronen, H. J., & Carlson, S. (2001). Working Memory of Identification of Emotional Vocal Expressions: An fMRI Study. *NeuroImage*, *13*(6), 1090–1101. https://doi.org/10.1006/nimg.2001.0777

Ramanan, S., & Bellana, B. (2019). A domain-general role for the angular gyrus in retrieving internal representations of the external world. *Journal of Neuroscience*, *39*(16), 2978–2980. https://doi.org/10.1523/JNEUROSCI.3231-18.2019

Ramanan, S., Piguet, O., & Irish, M. (2018). Rethinking the Role of the Angular Gyrus in Remembering the Past and Imagining the Future: The Contextual Integration Model. In *Neuroscientist* (Vol. 24, Issue 4, pp. 342–352). SAGE Publications Inc. https://doi.org/10.1177/1073858417735514

Ramsey, R. (2018). Neural Integration in Body Perception. *Journal of Cognitive Neuroscience*, *30*(10), 1442–1451. https://doi.org/10.1162/jocn_a_01299

Ramsey, R., & Ward, R. (2020). Putting the Nonsocial Into Social Neuroscience: A Role for Domain-General Priority Maps During Social Interactions. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691620904972

Ramsey, R., Hansen, P., Apperly, I., & Samson, D. (2013). Seeing it my way or your way: Frontoparietal brain areas sustain viewpoint-independent perspective selection processes. *Journal of Cognitive Neuroscience*, *25*(5), 670–684. https://doi.org/10.1162/jocn_a_00345

Rankin, K. P., Kramer, J. H., & Miller, B. L. (2005). Patterns of Cognitive and Emotional Empathy in Frontotemporal Lobar Degeneration. *Cognitive and Behavioral Neurology*, *18*(1), 28–36. https://doi.org/10.1097/01.wnn.0000152225.05377.ab

Raut, R. V., Snyder, A. Z., & Raichle, M. E. (2020). Hierarchical dynamics as a macroscopic organizing principle of the human brain. *Proceedings of the National Academy of Sciences*, *117*(34), 20890–20897. https://doi.org/10.1073/pnas.2003383117

Reilly, J., Hung, J., & Westbury, C. (2017). Non-Arbitrariness in Mapping Word Form to Meaning: Cross-Linguistic Formal Markers of Word Concreteness. *Cognitive Science*, *41*(4), 1071–1089. https://doi.org/10.1111/COGS.12361

Reilly, J., Peelle, J. E., Garcia, A., & Crutch, S. J. (2016). Linking somatic and symbolic representation in semantic memory: The dynamic multilevel reactivation framework. *Psychonomic Bulletin & Review 2015 23:4*, *23*(4), 1002–1014. https://doi.org/10.3758/S13423-015-0824-5

REFERENCES

Reilly, J., Westbury, C., Kean, J., & Peelle, J. E. (2012). Arbitrary Symbolism in Natural
Language Revisited: When Word Forms Carry Meaning. *PLOS ONE*, *7*(8), e42286.
https://doi.org/10.1371/JOURNAL.PONE.0042286

Relander, K., & Rämä, P. (2009). Separate neural processes for retrieval of voice identity and
word content in working memory. *Brain Research*, *1252*, 143–151.
https://doi.org/10.1016/j.brainres.2008.11.050

Reniers, R. L. E. P., Corcoran, R., Völlm, B. A., Mashru, A., Howard, R., & Liddle, P. F.
(2012). Moral decision-making, ToM, empathy and the default mode network.
Biological Psychology, *90*(3), 202–210.
https://doi.org/10.1016/j.biopsycho.2012.03.009

Rescorla, M. (2015). *The Computational Theory of Mind*.
https://seop.illc.uva.nl/entries/computational-mind/#EmbCog

Rice, G. E., Hoffman, P., & Lambon Ralph, M. A. (2015). Graded specialization within and
between the anterior temporal lobes. *Annals of the New York Academy of Sciences*,
1359(1), 84–97. https://doi.org/10.1111/nyas.12951

Rice, G. E., Hoffman, P., Binney, R. J., & Lambon Ralph, M. A. (2018). Concrete versus
abstract forms of social concept: An fMRI comparison of knowledge about people
versus social terms. *Philosophical Transactions of the Royal Society B: Biological
Sciences*, *373*(1752), 20170136. https://doi.org/10.1098/rstb.2017.0136

Rice, G. E., Lambon Ralph, M. A., & Hoffman, P. (2015). The Roles of Left Versus Right
Anterior Temporal Lobes in Conceptual Knowledge: An ALE Meta-analysis of 97
Functional Neuroimaging Studies. *Cerebral Cortex*, *25*, 4374–4391.
https://doi.org/10.1093/cercor/bhv024

Rilling, J. K., Dagenais, J. E., Goldsmith, D. R., Glenn, A. L., & Pagnoni, G. (2008). Social
cognitive neural networks during in-group and out-group interactions. *NeuroImage*,
41(4), 1447–1461. https://doi.org/10.1016/j.neuroimage.2008.03.044

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural
correlates of theory of mind within interpersonal interactions. *NeuroImage*, *22*(4),
1694–1703. https://doi.org/10.1016/j.neuroimage.2004.04.015

Rodd, J. M. (2020). Settling Into Semantic Space: An Ambiguity-Focused Account of Word-
Meaning Access. *Perspectives on Psychological Science*, *15*(2), 411–427.
https://doi.org/10.1177/1745691619885860

Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R.,
& Patterson, K. (2004). Structure and Deterioration of Semantic Memory: A

Neuropsychological and Computational Investigation. *Psychological Review*, *111*(1), 205–235. https://doi.org/10.1037/0033-295X.111.1.205

Rolls, E. T., Joliot, M., & Tzourio-Mazoyer, N. (2015). Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *NeuroImage*, *122*, 1–5. https://doi.org/10.1016/J.NEUROIMAGE.2015.07.075

Rosa, M. G. P., & Tweedale, R. (2005). Brain maps, great and small: Lessons from comparative studies of primate visual cortical organization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 665–691. https://doi.org/10.1098/RSTB.2005.1626

Rosema, S., Crowe, L., & Anderson, V. (2012). Social Function in Children and Adolescents after Traumatic Brain Injury: A Systematic Review 1989–2011. *Journal of Neurotrauma*, *29*(7), 1277–1291. https://doi.org/10.1089/neu.2011.2144

Rosen, H. J., Pace-Savitsky, K., Perry, R. J., Kramer, J. H., Miller, B. L., & Levenson, R. W. (2004). Recognition of Emotion in the Frontal and Temporal Variants of Frontotemporal Dementia. *Dementia and Geriatric Cognitive Disorders*, *17*(4), 277–281. https://doi.org/10.1159/000077154

Ross, L. A., & Olson, I. R. (2010). Social cognition and the anterior temporal lobes. *NeuroImage*, *49*(4), 3452–3462. https://doi.org/10.1016/j.neuroimage.2009.11.012

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

Rothmayr, C., Sodian, B., Hajak, G., Döhnel, K., Meinhardt, J., & Sommer, M. (2011). Common and distinct neural networks for false-belief reasoning and inhibitory control. *NeuroImage*, *56*(3), 1705–1713. https://doi.org/10.1016/j.neuroimage.2010.12.052

Roversi, C., Borghi, A. M., & Tummolini, L. (2013). A Marriage is an Artefact and not a Walk that We Take Together: An Experimental Study on the Categorization of Artefacts. *Review of Philosophy and Psychology*, *4*(3), 527–542. https://doi.org/10.1007/S13164-013-0150-7

Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). `Theory of mind' impairments and their relationship to executive functioning following frontal lobe excisions. *Brain*, *124*(3), 600–616. https://doi.org/10.1093/brain/124.3.600

RStudio Team. (2020). *RStudio: Integrated Development for R*. RStudio. https://www.rstudio.com/

Rushworth, M. F. S., Mars, R. B., & Sallet, J. (2013). Are there specialized circuits for social cognition and are they unique to humans? *Current Opinion in Neurobiology*, *23*(3), 436–442. https://doi.org/10.1016/J.CONB.2012.11.013

Saffran, E. M. (2000). The Organization of Semantic Memory: In Support of a Distributed Model. *Brain and Language*, *71*(1), 204–212. https://doi.org/10.1006/brln.1999.2251

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., & Smith, S. M. (2014). Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage*, *90*, 449–468. https://doi.org/10.1016/J.NEUROIMAGE.2013.11.046

Salo, T., Yarkoni, T., Nichols, T. E., Poline, J.-B., Bilgel, M., Bottenhorn, K. L., Jarecka, D., Kent, J. D., Kimbler, A., Nielson, D. M., Oudyk, K. M., Peraza, J. A., Pérez, A., Reeders, P. C., Yanes, J. A., & Laird, A. R. (2022). NiMARE: Neuroimaging Meta-Analysis Research Environment. *NeuroLibre*, *1*(1), 7. https://doi.org/10.55458/NEUROLIBRE.00007

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: A case of a selective deficit in inhibiting self-perspective. *Brain*, *128*(5), 1102–1111. https://doi.org/10.1093/brain/awh464

Samson, D., Houthuys, S., & Humphreys, G. W. (2015). Self-perspective inhibition deficits cannot be explained by general executive control difficulties. *Cortex*, *70*, 189–201. https://doi.org/10.1016/j.cortex.2014.12.021

Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: From imitation to Theory of Mind. *Cognition*, *122*(2), 228–235. https://doi.org/10.1016/j.cognition.2011.11.004

Satpute, A. B., & Lieberman, M. D. (2006). *Integrating automatic and controlled processes into neurocognitive models of social cognition*. https://doi.org/10.1016/j.brainres.2006.01.005

Satpute, A. B., Badre, D., & Ochsner, K. N. (2014). Distinct Regions of Prefrontal Cortex Are Associated with the Controlled Retrieval and Selection of Social Information. *Cerebral Cortex*, *24*(5), 1269–1277. https://doi.org/10.1093/cercor/bhs408

Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., Eickhoff, S. B., Hakonarson, H., Gur, R. C., Gur, R. E., & Wolf, D. H. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage*, *64*(1), 240–256. https://doi.org/10.1016/J.NEUROIMAGE.2012.08.052

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in 'theory of mind'. *NeuroImage*, *19*(4), 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Saxe, R., & Kanwisher, N. (2004). People Thinking about Thinking People: The Role of the Temporo-Parietal Junction in "Theory of Mind". In *Social Neuroscience*. Psychology Press.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, *43*(10), 1391–1399. https://doi.org/10.1016/j.neuropsychologia.2005.02.013

Saxe, R., Schulz, L. E., & Jiang, Y. V. (2006). Reading minds versus following rules: Dissociating theory of mind and executive control in the brain. *Social Neuroscience*, *1*(3–4), 284–298. https://doi.org/10.1080/17470910601000446

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, *19*(2), 65–72. https://doi.org/10.1016/j.tics.2014.11.007

Schenker, N. M., Buxhoeveden, D. P., Blackmon, W. L., Amunts, K., Zilles, K., & Semendeferi, K. (2008). A comparative quantitative analysis of cytoarchitecture and minicolumnar organization in Broca's area in humans and great apes. *Journal of Comparative Neurology*, *510*(1), 117–128. https://doi.org/10.1002/cne.21792

Schock, J., Cortese, M. J., & Khanna, M. M. (2012). Imageability estimates for 3,000 disyllabic words. *Behavior Research Methods*, *44*(2), 374–379. https://doi.org/10.3758/s13428-011-0162-0

Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E. N., & Saxe, R. (2009). Distinct Regions of Right Temporo-Parietal Junction Are Selective for Theory of Mind and Exogenous Attention. *PLOS ONE*, *4*(3), e4869. https://doi.org/10.1371/journal.pone.0004869

Schulte-Rüther, M., Markowitsch, H. J., Fink, G. R., & Piefke, M. (2007). Mirror Neuron and Theory of Mind Mechanisms Involved in Face-to-Face Interactions: A Functional Magnetic Resonance Imaging Approach to Empathy. *Journal of Cognitive Neuroscience*, *19*(8), 1354–1372. https://doi.org/10.1162/jocn.2007.19.8.1354

Schurz, M., Aichhorn, M., Martin, A., & Perner, J. (2013). Common brain areas engaged in false belief reasoning and visual perspective taking: A meta-analysis of functional brain imaging studies. *Frontiers in Human Neuroscience*, *7*, 712. https://doi.org/10.3389/fnhum.2013.00712

REFERENCES

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews*, *42*, 9–34. https://doi.org/10.1016/j.neubiorev.2014.01.009

Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., Sallet, J., & Kanske, P. (2020). Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin*, *147*(3), 293–327. https://doi.org/10.1037/bul0000303

Schurz, M., Tholen, M. G., Perner, J., Mars, R. B., & Sallet, J. (2017). Specifying the brain anatomy underlying temporo-parietal junction activations for theory of mind: A review using probabilistic atlases from different imaging modalities. *Human Brain Mapping*, *38*(9), 4788–4805. https://doi.org/10.1002/hbm.23675

Schuwerk, T., Schurz, M., Müller, F., Rupprecht, R., & Sommer, M. (2017). The rTPJ's overarching cognitive function in networks for attention and theory of mind. *Social Cognitive and Affective Neuroscience*, *12*(1), 157–168. https://doi.org/10.1093/scan/nsw163

Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2019). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*, *51*(3), 1258–1270. https://doi.org/10.3758/s13428-018-1099-3

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(3), 417–424. https://doi.org/10.1017/S0140525X00005756

Sepulcre, J., Sabuncu, M. R., Yeo, T. B., Liu, H., & Johnson, K. A. (2012). Stepwise Connectivity of the Modal Cortex Reveals the Multimodal Organization of the Human Brain. *The Journal of Neuroscience*, *32*(31), 10649–10661. https://doi.org/10.1523/JNEUROSCI.0759-12.2012

Shallice, T. (1987). Impairments of semantic processing: Multiple dissociations. In M. Coltheart, G. Sartoni, & R. Job (Eds.), *The Cognitive neuropsychology of language* (pp. 111–127). Lawrence Erlbaum Associates, Inc.

Shea, N. (2018). Metacognition and abstract concepts. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 373, Issue 1752). Royal Society Publishing. https://doi.org/10.1098/rstb.2017.0133

Sherman, J. W., Krieglmeyer, R., & Calanchini, J. (2014). Process Models Require Process Measures. In J. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 121–138). Guilford Press.

REFERENCES

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

Shimotake, A., Matsumoto, R., Ueno, T., Kunieda, T., Saito, S., Hoffman, P., Kikuchi, T., Fukuyama, H., Miyamoto, S., Takahashi, R., Ikeda, A., & Lambon Ralph, M. A. (2015). Direct Exploration of the Role of the Ventral Anterior Temporal Lobe in Semantic Memory: Cortical Stimulation and Local Field Potential Evidence From Subdural Grid Electrodes. *Cerebral Cortex*, *25*(10), 3802–3817. https://doi.org/10.1093/cercor/bhu262

Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., & Petersen, S. E. (1997). Common Blood Flow Changes across Visual Tasks: II. Decreases in Cerebral Cortex. *Journal of Cognitive Neuroscience*, *9*(5), 648–663. https://doi.org/10.1162/jocn.1997.9.5.648

Simmons, W. K., Reddish, M., Bellgowan, P. S. F., & Martin, A. (2010). The Selectivity and Functional Connectivity of the Anterior Temporal Lobes. *Cerebral Cortex*, *20*(4), 813–825. https://doi.org/10.1093/CERCOR/BHP149

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews*, *30*(6), 855–863. https://doi.org/10.1016/j.neubiorev.2006.06.011

Smallwood, J., Bernhardt, B. C., Leech, R., Bzdok, D., Jefferies, E., & Margulies, D. S. (2021). The default mode network in cognition: A topographical perspective. *Nature Reviews Neuroscience 2021 22:8*, *22*(8), 503–513. https://doi.org/10.1038/s41583-021-00474-4

Smallwood, J., Tipper, C., Brown, K., Baird, B., Engen, H., Michaels, J. R., Grafton, S., & Schooler, J. W. (2013). Escaping the here and now: Evidence for a role of the default mode network in perceptually decoupled thought. *NeuroImage*, *69*, 120–125. https://doi.org/10.1016/j.neuroimage.2012.12.012

Smith, R., Lane, R. D., Alkozei, A., Bao, J., Smith, C., Sanova, A., Nettles, M., & Killgore, W. D. S. (2017). Maintaining the feelings of others in working memory is associated with activation of the left anterior insula and left frontal-parietal control network. *Social Cognitive and Affective Neuroscience*, *12*(5), 848–860. https://doi.org/10.1093/scan/nsx011

Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D. A., Griffanti, L., Harms, M. P., Kelly, M., Laumann, T., Miller,

K. L., Moeller, S., Petersen, S., Power, J., Salimi-Khorshidi, G., Snyder, A. Z., Vu, A. T., … Glasser, M. F. (2013). Resting-state fMRI in the Human Connectome Project. *NeuroImage*, *80*, 144–168. https://doi.org/10.1016/j.neuroimage.2013.05.039

Snowden, J. S., Thompson, J. C., & Neary, D. (2012). Famous people knowledge and the right and left temporal lobes. *Behavioural Neurology*, *25*(1), 35–44. https://doi.org/10.3233/BEN-2012-0347

Snowden, J. S., Thompson, J. C., Neary, D., & Snowden, J. S. (2004). Knowledge of famous faces and names in semantic dementia. *Brain*, *127*, 860–872. https://doi.org/10.1093/brain/awh099

Snyder, H. R., Banich, M. T., & Munakata, Y. (2011). Choosing our words: Retrieval and selection processes recruit shared neural substrates in left ventrolateral prefrontal cortex. *Journal of Cognitive Neuroscience*, *23*(11), 3470–3482. https://doi.org/10.1162/jocn_a_00023

Soch, J., Deserno, L., Assmann, A., Barman, A., Walter, H., Richardson-Klavehn, A., & Schott, B. H. (2017). Inhibition of information flow to the default mode network during self-reference versus reference to others. *Cerebral Cortex*, *27*(8), 3930–3942. https://doi.org/10.1093/cercor/bhw206

Soni, M., Lambon Ralph, M. A., Noonan, K., Ehsan, S., Hodgson, C., & Woollams, A. M. (2009). "L" is for tiger: Effects of phonological (mis)cueing on picture naming in semantic aphasia. *Journal of Neurolinguistics*, *22*(6), 538–547. https://doi.org/10.1016/j.jneuroling.2009.06.002

Souter, N. E., Lindquist, K. A., & Jefferies, E. (2021). Impaired emotion perception and categorization in semantic aphasia. *Neuropsychologia*, *162*, 108052. https://doi.org/10.1016/j.neuropsychologia.2021.108052

Sowden, S., & Shah, P. (2014). Self-other control: A candidate mechanism for social cognitive function. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00789

Spisák, T., Spisák, Z., Zunhammer, M., Bingel, U., Smith, S., Nichols, T., & Kincses, T. (2019). Probabilistic TFCE: A generalized combination of cluster size and voxel intensity to increase statistical power. *NeuroImage*, *185*, 12–26. https://doi.org/10.1016/J.NEUROIMAGE.2018.09.078

Spreng, R. N. (2013). Examining the role of memory in social cognition. *Frontiers in Psychology*, *4*. https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00437

REFERENCES

Spreng, R. N., & Andrews-Hanna, J. R. (2015). The Default Network and Social Cognition. In A. W. Toga (Ed.), *Brain Mapping: An Encyclopedic Reference* (pp. 165–169). Academic press. https://doi.org/10.1016/B978-0-12-397025-1.00173-1

Spreng, R. N., DuPre, E., Selarka, D., Garcia, J., Gojkovic, S., Mildner, J., Luh, W.-M., & Turner, G. R. (2014). Goal-Congruent Default Network Activity Facilitates Cognitive Control. *Journal of Neuroscience*, *34*(42), 14108–14114. https://doi.org/10.1523/JNEUROSCI.2815-14.2014

Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W., & Schacter, D. L. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *NeuroImage*, *53*(1), 303–317. https://doi.org/10.1016/j.neuroimage.2010.06.016

Spunt, R. P., & Adolphs, R. (2014). Validating the Why/How contrast for functional MRI studies of Theory of Mind. *NeuroImage*, *99*, 301–311. https://doi.org/10.1016/j.neuroimage.2014.05.023

Spunt, R. P., & Adolphs, R. (2017). A new look at domain specificity: Insights from social neuroscience. *Nature Reviews Neuroscience*, *18*(9), 559–567. https://doi.org/10.1038/nrn.2017.76

Spunt, R. P., & Lieberman, M. D. (2012). An integrative model of the neural systems supporting the comprehension of observed emotional behavior. *NeuroImage*, *59*(3), 3050–3059. https://doi.org/10.1016/j.neuroimage.2011.10.005

Spunt, R. P., Satpute, A. B., & Lieberman, M. D. (2010). Identifying the What, Why, and How of an Observed Action: An fMRI Study of Mentalizing and Mechanizing during Action Observation. *Journal of Cognitive Neuroscience*, *23*(1), 63–74. https://doi.org/10.1162/jocn.2010.21446

Steinbeis, N. (2016). The role of self–other distinction in understanding others' mental and emotional states: Neurocognitive mechanisms in children and adults. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150074. https://doi.org/10.1098/rstb.2015.0074

Stevens, M., & Brysbaert, M. (2016). *A simple solution for missing observations based on random effects models*.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal Lobe Contributions to Theory of Mind. *Journal of Cognitive Neuroscience*, *10*(5), 640–656. https://doi.org/10.1162/089892998562942

REFERENCES

Strik-Lievers, F., Bolognesi, M., & Winter, B. (2021). The linguistic dimensions of concrete and abstract concepts: Lexical category, morphological structure, countability, and etymology. *Cognitive Linguistics*, *32*(4), 641–670. https://doi.org/10.1515/COG-2021-0007/ASSET/GRAPHIC/J_COG-2021-0007_FIG_005.JPG

Stuss, D. T., Gallup, G. G., Jr, & Alexander, M. P. (2001). The frontal lobes are necessary for `theory of mind'. *Brain*, *124*(2), 279–286. https://doi.org/10.1093/brain/124.2.279

Suárez, L. E., Markello, R. D., Betzel, R. F., & Misic, B. (2020). Linking Structure and Function in Macroscale Brain Networks. *Trends in Cognitive Sciences*, *24*(4), 302–315. https://doi.org/10.1016/J.TICS.2020.01.008

Swick, D., Ashley, V., & Turken, A. U. (2008). Left inferior frontal gyrus is critical for response inhibition. *BMC Neuroscience*, *9*(1), 102. https://doi.org/10.1186/1471-2202-9-102

Tamm, S., Nilsonne, G., Schwarz, J., Lamm, C., Kecklund, G., Petrovic, P., Fischer, H., Åkerstedt, T., & Lekander, M. (2017). The effect of sleep restriction on empathy for pain: An fMRI study in younger and older adults. *Scientific Reports*, *7*(1), 1–14. https://doi.org/10.1038/s41598-017-12098-9

Taylor, P., Hobbs, J. N., Burroni, J., & Siegelmann, H. T. (2015). The global landscape of cognition: Hierarchical aggregation as an organizational principle of human cortical networks and functions. *Scientific Reports*, *5*(1), Article 1. https://doi.org/10.1038/srep18112

Thiebaut de Schotten, M., Urbanski, M., Batrancourt, B., Levy, R., Dubois, B., Cerliani, L., & Volle, E. (2017). Rostro-caudal Architecture of the Frontal Lobes in Humans. *Cerebral Cortex*, *27*(8), 4033–4047. https://doi.org/10.1093/cercor/bhw215

Thioux, M., Gazzola, V., & Keysers, C. (2008). Action Understanding: How, What and Why. *Current Biology*, *18*(10), R431–R434. https://doi.org/10.1016/j.cub.2008.03.018

Thomas Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fisch, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. https://doi.org/10.1152/jn.00338.2011

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(26), 14792–14797. https://doi.org/10.1073/pnas.94.26.14792

REFERENCES

Thompson, B., Roberts, S. G., & Lupyan, G. (2020). Cultural influences on word meanings revealed through large-scale semantic alignment. *Nature Human Behaviour 2020 4:10*, *4*(10), 1029–1038. https://doi.org/10.1038/s41562-020-0924-8

Thompson, E. L., Bird, G., & Catmur, C. (2019). Conceptualizing and testing action understanding. *Neuroscience & Biobehavioral Reviews*, *105*, 106–114. https://doi.org/10.1016/j.neubiorev.2019.08.002

Thompson, H. E., Almaghyuli, A., Noonan, K. A., Barak, O., Lambon Ralph, M. A., & Jefferies, E. (2018). The contribution of executive control to semantic cognition: Convergent evidence from semantic aphasia and executive dysfunction. *Journal of Neuropsychology*, *12*(2), 312–340. https://doi.org/10.1111/jnp.12142

Thompson, S. A., Patterson, K., & Hodges, J. R. (2003). Left/right asymmetry of atrophy in semantic dementia. *Neurology*, *61*(9), 1196 LP – 1203. https://doi.org/10.1212/01.WNL.0000091868.28557.B8

Tian, Y., & Zalesky, A. (2018). Characterizing the functional connectivity diversity of the insula cortex: Subregions, diversity curves and behavior. *NeuroImage*, *183*, 716–733. https://doi.org/10.1016/j.neuroimage.2018.08.055

Timmers, I., Park, A. L., Fischer, M. D., Kronman, C. A., Heathcote, L. C., Hernandez, J. M., & Simons, L. E. (2018). Is Empathy for Pain Unique in Its Neural Correlates? A Meta-Analysis of Neuroimaging Studies of Empathy. *Frontiers in Behavioral Neuroscience*, *12*, 289. https://doi.org/10.3389/fnbeh.2018.00289

Tomasello, M. (2009). *Why we cooperate*. MIT Press.

Tomasello, M. (2020). The adaptive origins of uniquely human sociality. *Philosophical Transactions of the Royal Society B*, *375*(1803). https://doi.org/10.1098/RSTB.2019.0493

Tops, M., & Boksem, M. (2011). A Potential Role of the Inferior Frontal Gyrus and Anterior Insula in Cognitive Control, Brain Rhythms, and Event-Related Potentials. *Frontiers in Psychology*, *2*. https://www.frontiersin.org/articles/10.3389/fpsyg.2011.00330

Troche, J., Crutch, S. J., & Reilly, J. (2017). Defining a conceptual topography of word concreteness: Clustering properties of emotion, sensation, and magnitude among 750 english words. *Frontiers in Psychology*, *8*, 1787. https://doi.org/10.3389/fpsyg.2017.01787

Troche, J., Crutch, S., & Reilly, J. (2014). Clustering, hierarchical organization, and the topography of abstract and concrete nouns. *Frontiers in Psychology*, *5*, 360. https://doi.org/10.3389/FPSYG.2014.00360

REFERENCES

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. *Human Brain Mapping*, *33*(1), 1–13. https://doi.org/10.1002/hbm.21186

Uddén, J., & Bahlmann, J. (2012). A rostro-caudal gradient of structured sequence processing in the left inferior frontal gyrus. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 2023–2032. https://doi.org/10.1098/rstb.2012.0009

Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature Reviews Neuroscience*, *16*(1), Article 1. https://doi.org/10.1038/nrn3857

Uddin, L. Q., Betzel, R. F., Cohen, J. R., Damoiseaux, J. S., Brigard, F. D., Eickhoff, S. B., Fornito, A., Gratton, C., Gordon, E. M., Laird, A., Larson-Prior, L. J., McIntosh, A. R., Nickerson, L. D., Pessoa, L., Pinho, A. L., Poldrack, R., Razi, A., Sadaghiani, S., Shine, J. M., … Spreng, R. N. (2022). *Controversies and current progress on large-scale brain network nomenclature from OHBM WHATNET: Workgroup for HArmonized Taxonomy of NETworks*. OSF Preprints. https://doi.org/10.31219/osf.io/25za6

Ueno, T., Saito, S., Rogers, T. T., & Lambon Ralph, M. A. (2011). Lichtheim 2: Synthesizing Aphasia and the Neural Basis of Language in a Neurocomputational Model of the Dual Dorsal-Ventral Language Pathways. *Neuron*, *72*(2), 385–396. https://doi.org/10.1016/j.neuron.2011.09.013

Uleman, J. S., Saribay, S. A., & Gonzalez, C. M. (2007). Spontaneous Inferences, Implicit Impressions, and Implicit Theories. *Annual Review of Psychology*, *59*, 329–360. https://doi.org/10.1146/annurev.psych.59.103006.093707

Van der Meer, L., Groenewold, N. A., Nolen, W. A., Pijnenborg, M., & Aleman, A. (2011). Inhibit yourself and understand the other: Neural basis of distinct processes underlying Theory of Mind. *NeuroImage*, *56*(4), 2364–2374. https://doi.org/10.1016/j.neuroimage.2011.03.053

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage*, *80*, 62–79. https://doi.org/10.1016/J.NEUROIMAGE.2013.05.041

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, *30*(3), 829–858). https://doi.org/10.1002/hbm.20547

REFERENCES

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage*, *48*(3), 564–584. https://doi.org/10.1016/j.neuroimage.2009.06.009

Van Overwalle, F., & Vandekerckhove, M. (2013). Implicit and explicit social mentalizing: Dual processes driven by a shared neural network. *Frontiers in Human Neuroscience*, *7*(SEP). https://doi.org/10.3389/fnhum.2013.00560

Vargas, R., & Just, M. A. (2020). Neural Representations of Abstract Concepts: Identifying Underlying Neurosemantic Dimensions. *Cerebral Cortex*, *30*(4), 2157–2166. https://doi.org/10.1093/CERCOR/BHZ229

Vatansever, D., Smallwood, J., & Jefferies, E. (2021). Varying demands for cognitive control reveals shared neural processes supporting semantic and episodic memory retrieval. *Nature Communications*, *12*(1), 2134. https://doi.org/10.1038/s41467-021-22443-2

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*(4), 422–488. https://doi.org/10.1016/J.COGPSYCH.2003.09.001

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, *30*(4), 1414–1432. https://doi.org/10.1016/j.neuroimage.2005.11.002

Villani, C., D'Ascenzo, S., Borghi, A. M., Roversi, C., Benassi, M., & Lugli, L. (2021). Is justice grounded? How expertise shapes conceptual representation of institutional concepts. *Psychological Research*, *1*, 1–17. https://doi.org/10.1007/S00426-021-01492-8

Villani, C., Lugli, L., Liuzza, M., & Borghi, A. M. (2019). Varieties of abstract concepts and their multiple dimensions. *Language and Cognition*, *11*(3), 403–430. https://doi.org/10.1017/LANGCOG.2019.23

Visser, M., & Lambon Ralph, M. A. (2011). Differential Contributions of Bilateral Ventral Anterior Temporal Lobe and Left Anterior Superior Temporal Gyrus to Semantic Processes. *Journal of Cognitive Neuroscience*, *23*(10), 3121–3131. https://doi.org/10.1162/jocn_a_00007

Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic processing in the anterior temporal lobes: A meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, *22*(6), 1083–1094. https://doi.org/10.1162/jocn.2009.21309

REFERENCES

Visser, M., Jefferies, E., Embleton, K. V., & Lambon Ralph, M. A. (2012). Both the Middle
Temporal Gyrus and the Ventral Anterior Temporal Area Are Crucial for Multimodal
Semantic Processing: Distortion-corrected fMRI Evidence for a Double Gradient of
Information Convergence in the Temporal Lobes. *Journal of Cognitive Neuroscience*,
24(8), 1766–1778. https://doi.org/10.1162/jocn_a_00244

Vos de Wael, R., Benkarim, O., Paquola, C., Lariviere, S., Royer, J., Tavakol, S., Xu, T.,
Hong, S. J., Langs, G., Valk, S., Misic, B., Milham, M., Margulies, D., Smallwood, J.,
& Bernhardt, B. C. (2020). BrainSpace: A toolbox for the analysis of macroscale
gradients in neuroimaging and connectomics datasets. *Communications Biology*, *3*(1),
1–10. https://doi.org/10.1038/s42003-020-0794-7

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning:
Left prefrontal cortex guides controlled semantic retrieval. *Neuron*, *31*(2), 329–338.
https://doi.org/10.1016/S0896-6273(01)00359-2

Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2006). Developmental changes in the
neural basis of interpreting communicative intent. *Social Cognitive and Affective
Neuroscience*, *1*(2), 107–121. https://doi.org/10.1093/scan/nsl018

Wang, J.X., Li, Y., Mu, Y., & Zhuang, J.Y. (2022). Common and unique neural mechanisms
of social and nonsocial conflict resolving and adaptation. *Cerebral Cortex*, bhac306.
https://doi.org/10.1093/cercor/bhac306

Wang, J., Yang, Y., Zhao, X., Zuo, Z., & Tan, L. H. (2020). Evolutional and developmental
anatomical architecture of the left inferior frontal gyrus. *NeuroImage*, *222*, 117268.
https://doi.org/10.1016/j.neuroimage.2020.117268

Wang, X., Bernhardt, B. C., Karapanagiotidis, T., De Caso, I., Gonzalez Alam, T. R. del J.,
Cotter, Z., Smallwood, J., & Jefferies, E. (2018). The structural basis of semantic
control: Evidence from individual differences in cortical thickness. *NeuroImage*, *181*,
480–489. https://doi.org/10.1016/J.NEUROIMAGE.2018.07.044

Wang, X., Margulies, D. S., Smallwood, J., & Jefferies, E. (2020). A gradient from long-term
memory to novel cognition: Transitions through default mode and executive cortex.
NeuroImage, *220*. https://doi.org/10.1101/2020.01.16.908327

Wang, X., Wang, B., & Bi, Y. (2019). Close yet independent: Dissociation of social from
valence and abstract semantic dimensions in the left anterior temporal lobe. *Human
Brain Mapping*, *40*(16), 4759–4776. https://doi.org/10.1002/HBM.24735

Wang, Y., Collins, J. A., Koski, J., Nugiel, T., Metoki, A., & Olson, I. R. (2017). Dynamic
neural architecture for social knowledge retrieval. *Proceedings of the National*

Academy of Sciences, *114*(16), E3305–E3314.
https://doi.org/10.1073/pnas.1621234114

Warnell, K. R., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition*, *191*, 103997. https://doi.org/10.1016/j.cognition.2019.06.009

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Warrington, E. K., & McCarthy, R. A. (1994). Multiple meaning systems in the brain: A case for visual semantics. *Neuropsychologia*, *32*(12), 1465–1473. https://doi.org/10.1016/0028-3932(94)90118-X

Weissman, D. H., Perkins, A. S., & Woldorff, M. G. (2008). Cognitive control in social situations: A role for the dorsolateral prefrontal cortex. *NeuroImage*, *40*(2), 955–962. https://doi.org/10.1016/j.neuroimage.2007.12.021

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2011). The Neural Organization of Semantic Control: TMS Evidence for a Distributed Network in Left Inferior Frontal and Posterior Middle Temporal Gyrus. *Cerebral Cortex*, *21*(5), 1066–1075. https://doi.org/10.1093/CERCOR/BHQ180

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2012). Executive semantic processing is underpinned by a large-scale neural network: Revealing the contribution of left prefrontal, posterior temporal, and parietal cortex to controlled retrieval and selection using TMS. *Journal of Cognitive Neuroscience*, *24*(1), 133–147. https://doi.org/10.1162/jocn_a_00123

Wiemer-Hastings, K. K., & Xu, X. (2005). Content Differences for Abstract and Concrete Concepts. *Cognitive Science*, *29*(5), 719–736. https://doi.org/10.1207/S15516709COG0000_33

Wilson-Mendenhall, C. D., Simmons, W. K., Martin, A., & Barsalou, L. W. (2013). Contextual processing of abstract concepts reveals neural representations of nonlinguistic semantic content. *Journal of Cognitive Neuroscience*, *25*(6), 920–935. https://doi.org/10.1162/JOCN_A_00361

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

REFERENCES

Wojtasik, M., Bludau, S., Eickhoff, S. B., Mohlberg, H., Gerboga, F., Caspers, S., & Amunts, K. (2020). Cytoarchitectonic Characterization and Functional Decoding of Four New Areas in the Human Lateral Orbitofrontal Cortex. *Frontiers in Neuroanatomy*, *14*, 2. https://doi.org/10.3389/FNANA.2020.00002/BIBTEX

Wurm, M. F., & Lingnau, A. (2015). Decoding Actions at Different Levels of Abstraction. *Journal of Neuroscience*, *35*(20), 7727–7735. https://doi.org/10.1523/JNEUROSCI.0188-15.2015

Xin, F., & Lei, X. (2015). Competition between frontoparietal control and default networks supports social working memory and empathy. *Social Cognitive and Affective Neuroscience*, *10*(8), 1144–1152. https://doi.org/10.1093/scan/nsu160

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N : A new measure of orthographic similarity. *Psychonomic Bulletin & Review 2008 15:5*, *15*(5), 971–979. https://doi.org/10.3758/PBR.15.5.971

Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review 2015 23:4*, *23*(4), 1015–1027. https://doi.org/10.3758/S13423-015-0948-7

Young, L., & Saxe, R. (2009). An fMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Neuroscience*, *21*(7), 1396–1405. https://doi.org/10.1162/jocn.2009.21137

Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences*, *107*(15), 6753–6758. https://doi.org/10.1073/pnas.0914826107

Young, L., Scholz, J., & Saxe, R. (2011). Neural evidence for 'intuitive prosecution': The use of mental state information for negative moral verdicts. *Social Neuroscience*, *6*(3), 302–315. https://doi.org/10.1080/17470919.2010.529712

Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(15), 6430–6435. https://doi.org/10.1073/pnas.0607061104

Zaki, J. (2013). Cue Integration: A Common Framework for Social Cognition and Physical Perception. *Perspectives on Psychological Science*, *8*(3), 296–312. https://doi.org/10.1177/1745691613475454

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

REFERENCES

Journal of Neuroscience, *30*(25), 8481–8488.

https://doi.org/10.1523/JNEUROSCI.0382-10.2010

Zdrazilova, L., & Pexman, P. M. (2013). Grasping the invisible: Semantic processing of

abstract words. *Psychonomic Bulletin and Review*, *20*(6), 1312–1318.

https://doi.org/10.3758/s13423-013-0452-x

Zhang, G., Xu, Y., Zhang, M., Wang, S., & Lin, N. (2021). The brain network in support of

social semantic accumulation. *Social Cognitive and Affective Neuroscience*, *16*(4),

393–405. https://doi.org/10.1093/SCAN/NSAB003

APPENDIX I. Supplementary Materials

for

CHAPTER 3. Quantifying social semantics: an inclusive definition of socialness and ratings for 8,388 English words

Section S1. Pilot Study

1.1. Participants

Before starting the main experiment, we tested our socialness rating task in a sample of 36 participants (23 female, 13 male; M_{age} = 22.94 years, SD_{age} = 6.5). Participants were recruited from the participant pool at Bangor University. Participants completed the rating task in 24 minutes on average and were compensated with course credit. Of the participants, 17 saw version 1 of the instructions and 19 saw version 2.

1.2. Materials

We selected 60 items (including nouns, adjectives and verbs) that span the following dimensions: valence (Warriner et al., 2013), concreteness (Brysbaert et al., 2014) and social interaction (Binder et al., 2016; Troche et al., 2017). We created two versions of the instructions to assess whether wording influenced participants' understanding of the instructions and their ratings. In version 1, socialness was defined as the degree to which a word's meaning has a social quality, whereas in version 2 it was defined as the degree to which a word's meaning has social relevance. The rest of the instructions and examples were identical in the two versions.

1.3. Procedure

The word stimuli were presented using Qualtrics and linked to the online participant recruitment platform at Bangor University. Following the consent form, demographics survey and instructions, participants rated how well they understood the instructions on a 5-point Likert scale from not at all (1) to extremely well (5). Then, participants proceeded to rate all the items using a 7-point Likert scale presented horizontally below each word. In addition,

there was an "I don't know the meaning of this word" option. Items were presented in random order and only one word was presented per page. Following 30% of the items, participants were asked to explain the reasoning behind their chosen rating for the respective word. At the end of the ratings task, participants were asked to explain the task instructions using a text box.

1.4. Results

1.4.1. Understanding of instructions

The data and analysis scripts can be accessed via the OSF project page (https://osf.io/2dqnj/). 8.33% of the pilot participant sample reported understanding the instructions moderately well, 55.56% very well and 36.11% extremely well. Most participants provided explanations that were consistent with our inclusive socialness definition (e.g.: "*Socialness describes anything related to people, for instance their interactions, ways of describing people, relationships, social places/event, social beliefs, etc.*"; "*relation to people and society*"). Interestingly, many participants focused their explanations on the link between the word's meaning and social interactions (e.g., "*Of, or relating to, the interaction of individuals. That which pertains to people interacting.*"; "*The socialness referred to how likely it was that the words were associated with social concepts by involving social interactions, by influencing social interactions, and by representing those and the values of those in social encounters.*").

1.4.2. Reliability

We examined the reliability of the ratings by computing the split half reliability for the 60 words. We found a mean Spearman-Brown corrected split-half reliability of 0.97 ($SD = 0.12$) across 100 random splits, suggesting high reliability. In addition, we assessed inter-rater reliability by computing the two-way random-effects intra-class correlation coefficient (ICC) based on absolute agreement. We found an ICC(2,1) = 0.4, 95%CI [0.33, 0.48] suggesting poor to moderate reliability of individual ratings and an ICC(2, 36) = 0.96, 95%CI [0.95, 0.97] suggesting excellent reliability of the average ratings across 36 raters. Moreover, we found an ICC of 0.94, 95%CI [0.92, 0.96] which suggested that the 19 raters who saw the instructions eventually used in the main experiment (version 2) provided average ratings that were highly reliable.

1.4.3. The influence of instruction version

A Pearson's Chi-square test of independence suggested that self-reported understanding of the instructions did not depend on the version of the instructions χ^2 (2, N = 36) = 0.92, p = 0.63. There was a strong positive correlation between mean socialness scores for the two instruction versions (r = 0.93, p <.001, R^2 = 0.87). Moreover, the reliability was comparable for the two versions, with a mean Spearman-Brown corrected split-half reliability of 0.93 (SD = 0.15) for version 1 and 0.95 (SD = 0.12) for version 2 (across 100 random splits). Therefore, we concluded that the wording did not significantly influence raters' responses.

Section S2. Rating task instructions

Our society and interactions with other people feature at the heart of many of our experiences in life and this study explores whether this is reflected in the language we use.

In particular, this study is interested in the 'socialness' of words, or the degree to which words can be considered 'social'. There are many ways in which a word can be considered 'social'. For example, a word is considered to be 'social' if it describes or refers to a social characteristic of a person or group of people (e.g., 'trustworthy'), a social behaviour or interaction (e.g., ' to fight'), a social role (e.g., 'teacher'), a social space (e.g., 'pub'), a social institution (e.g., 'hospital') or system (e.g., 'nation'), a social value (e.g., 'righteousness') or ideology (e.g., 'feminism'), or any other socially-relevant concept. In contrast, 'non-social' words have meanings that lack in social relevance (e.g., 'chair', 'time').
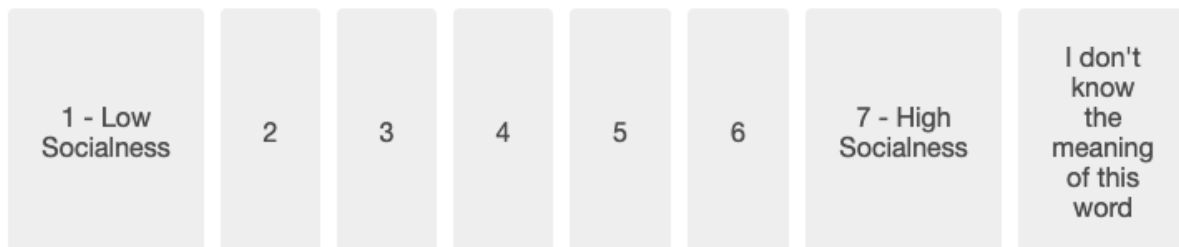
Words also differ in the degree to which they can be considered social. Some words clearly refer to social things, social qualities or to social actions or events (e.g., 'trustworthy') whereas, for other words, the relationship to society or social interactions might only become apparent after a period of intense thought (e.g., 'promotion'), or not at all (e.g., 'chair'). The purpose of this study is to rate words based on the degree to which they have a meaning that has social relevance. Any word that in your estimation refers to something that has clear social relevance should be given a **high 'socialness' rating** (at the upper end of the numerical scale). Any word that in your estimation refers to something lacking in social relevance should be given a **low 'socialness' rating** (at the lower end of the scale). Any word that in your estimation refers to

something that is not fundamentally social but has some social elements (e.g., 'smartphone'), or can be thought of as social in some circumstances (e.g., 'event'), should be given an intermediate socialness rating.

It is important that you base these ratings on the degree of the social relevance of the word's meaning and not whether the meaning is prosocial versus antisocial or evokes positive/negative associations. For example, the word 'fight' should be given a high 'socialness' rating because it refers to a type of interaction between people, and even though the interaction is antisocial.

Please make your 'socialness' ratings using the 7-point scale. A value of 1 indicates a low 'socialness' rating, and a value of 7 indicates a high 'socialness' rating. Values of 2 to 6 indicate intermediate ratings. Please feel free to use the whole range of values provided when making your ratings. Click on the rating that is most appropriate for each word. When making your ratings, try to be as accurate as possible, but do not spend too much time on any one word. If you are not familiar with a word's meaning, please select 'I don't know the meaning of this word'.

| 1 - Low Socialness | 2 | 3 | 4 | 5 | 6 | 7 - High Socialness | I don't know the meaning of this word |

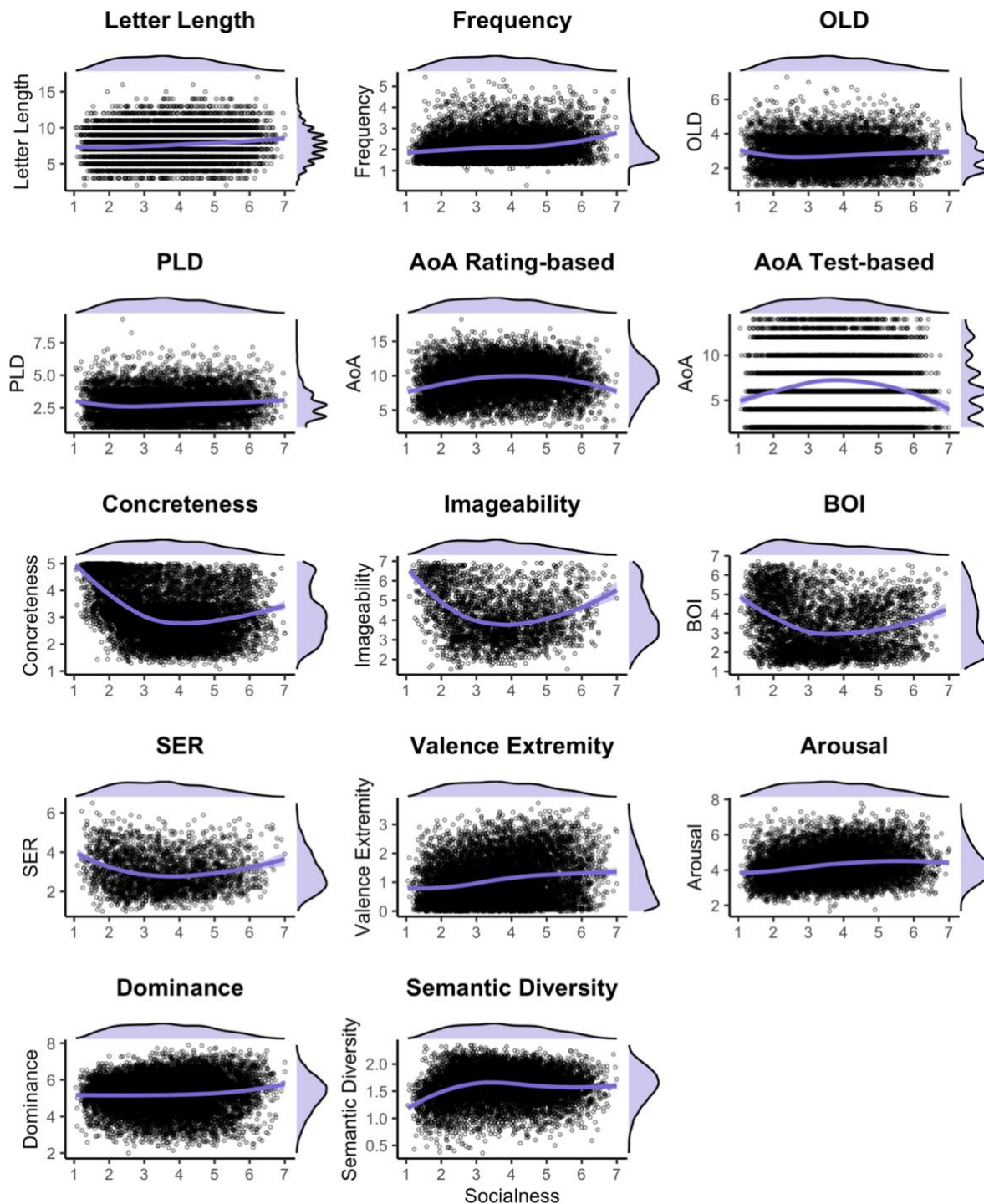### Section S3. Additional results

Figure S1. Scatterplots illustrate the relationships (highlighted by the loess line) between mean socialness ratings (x axis) and lexical-semantic dimensions (y axis). The density distributions of the socialness and variables of interest are plotted on the top and right of the graphs respectively. For each variable of interest, the numbers of items in common with our socialness ratings are as follows: length, concreteness, valence, arousal and dominance: 8,388; log subtitle frequency: 8,160; OLD and PLD: 8,027; rating-based AoA: 8,348; test-based AoA: 7, 321; imageability: 2,680; BOI: 4,038;

SER: 2,645. SER = sensory experience rating; BOI = body-object interaction; AoA = age of acquisition; PLD = phonologic Levenshtein distance; OLD = orthographic Levenshtein distance.

Table S1. *Means, standard deviations and correlations of all variables of interest for the regression analysis predicting performance in the English Lexicon Project Lexical Decision Task (N = 6,926)*

Variable	M	SD	1	2	3	4	5	6	7	8
1. Length	7.54	2.13								
2. Frequency	2.18	0.65	-.36**							
3. Age of Acquisition	9.43	2.44	.31**	-.57**						
4. Socialness	3.67	1.24	.13**	.18**	.09**					
5. Concreteness	3.08	0.95	-.06**	.07**	-.35**	-.29**				
6. Valence Extremity	1.07	0.77	.01	.14**	-.15**	.23**	-.14**			
7. Semantic Diversity	1.60	0.31	-.13**	.34**	-.16**	.09**	-.40**	.03*		
8. LDT zRT	-0.25	0.31	.52**	-.59**	.56**	-.03*	-.11**	-.08**	-.25**	
9. LDT Error Rate	0.06	0.08	-.08**	-.33**	.37**	-.10**	-.06**	-.10**	-.15**	.51**

Note. M and SD are used to represent mean and standard deviation, respectively. LDT = lexical decision task; zRT = standardized reaction times. * indicates p < .05. ** indicates p < .01.

Table S2. Means, standard deviations and correlations of all variables of interest for the regression analysis predicting performance in the English Crowdsourcing Project Word Knowledge Task (N = 7,010).

Variable	M	SD	1	2	3	4	5	6	7	8
1. Length	7.55	2.13								
2. Frequency	2.17	0.65	-.36**							
3. Age of Acquisition	9.43	2.43	.31**	-.57**						
4. Socialness	3.67	1.24	.13**	.18**	.09**					
5. Concreteness	3.08	0.95	-.06**	.07**	-.35**	-.29**				
6. Valence Extremity	1.07	0.77	.01	.14**	-.14**	.23**	-.14**			
7. Semantic Diversity	1.59	0.31	-.13**	.34**	-.16**	.09**	-.40**	.03*		
8. Recognition zRT	-0.53	0.12	.39**	-.55**	.53**	-.05**	-.11**	-.14**	-.23**	
9. Proportion Unknown	0.01	0.02	-.06**	-.35**	.38**	-.07**	-.06**	-.13**	-.18**	.63**

Note. M and SD are used to represent mean and standard deviation, respectively. zRT = standardized reaction times. * indicates p < .05. ** indicates p < .01.

APPENDIX II. Supplementary Materials

for

CHAPTER 4. Establishing a role of the semantic control network in social cognitive processing: a meta-analysis of functional neuroimaging studies

Supplementary Information No. 2 (list of included experiments) and No. 3 (results tables) can be accessed at: https://osf.io/fktb8/

Supplementary Information No. 1 (SI1)

Section S1. Literature search and data extraction

S.1.1. Deviations from the pre-registration

We pre-registered our plans for the current study after accessing some of the raw data and conducting independent ALE analyses on subsets of the final semantic control, theory of mind and trait inference datasets (see the OSF pre-registration for further details: https://osf.io/dscwv). There are two deviations from our pre-registered protocols which are detailed below:

1. *Inclusion/exclusion criteria*. We pre-registered the exclusion of contrasts between experimental conditions and low-level baselines (e.g., rest, fixation cross). However, we subsequently decided to collect and include data from contrasts against low-level baselines because they can reveal activity associated with domain-general cognitive processes that is subtracted out by contrasts between active conditions. This could include semantic and executive processes that are common to both social and non-social tasks and which were of interest for testing our pre-registered hypotheses. For completeness, we report the analyses whilst excluding this subset of contrasts on the project's OSF page (osf.io/fktb8/).

2. *Data collection*. In our pre-registration, empathy was considered a single sub-domain of social cognition. However, the relevant literature distinguishes between empathy for affective states and empathy for pain because there are important differences in the brain regions underpinning these two types of empathic processing (Ding et al., 2020; Kogler et al., 2020; Timmers et al., 2018). Thus, pooling across empathy for pain and emotions would preclude the identification of all brain regions engaged in empathy. Therefore, we made a distinction between experiments looking at empathy for pain and those investigating empathy for emotions, and conducted all analyses separately for these two subsets. For completeness, the results of a global empathy (pooled across pain and emotions) are reported in Section S3.

S.1.2. Comparison between the present meta-analyses and the prior meta-analyses

Table S1. A summary of the similarities and differences in the methodological approach taken in the present meta-analyses and the prior meta-analyses upon which updates were based.

	Semantic Control Noonan et al. (2013)	Theory of Mind Molenberghs et al. (2016)	Empathy Timmers et al., (2018)	Moral Reasoning Eres et al., (2018)
Task definition		c		
Contrast definition		d		e
Included results from both fMRI and PET studies				
Included whole-brain analyses only				
Included multiple contrasts from a single sample				
Adjusted for multiple contrasts from a single sample	a	b	a	a
Employed an updated GingerALE algorithm thereby avoiding overly liberal statistical thresholds (g)		a		
Employed recommended FWE cluster-extent correction (g)	f	f	f	f
Employed additional conservative height-based thresholding				

Orange coloured cells indicate differences whereas green cells indicate an identical or equivalent approach. FWE = Family wise error

[a]Not reported;

[b]Not applicable;

[c] Unlike the present study, Molenberghs et al. (2016) included studies of irony comprehension, trait inference & those employing 'interactive games' (see Section 2.1.2);

[d] Unlike the present study, Molenberghs et al. (2016) did not include contrasts between an experimental social condition and low-level (e.g., fixation) baseline conditions (see Section 2.1.2);

[e] Unlike the present study, Eres et al., (2018) included contrasts between different components of moral reasoning (e.g., utilitarian vs. deontological moral judgements) and did not include contrasts between experimental conditions and low-level baseline conditions(see Section 2.1.2);

[f] Used False Discovery Rate correction.

[g](Eickhoff et al., 2016)

S.1.3. Data extraction

We extracted the following information from each study and included it in our database: authors, year of publication, sample size, imaging method (fMRI/PET), task description, contrast, type of contrast (low-level baseline vs high-level baseline), coordinate space (MNI/Talairach), type of instructional cue (explicit/implicit) and type of difficulty category based on participant reaction times (experimental condition harder/easier than or equally difficult to the control condition). In addition, we collected information about other potential moderating factors not investigated in this study as follows:

- Semantic control: type of cognitive manipulation (semantic production/ semantic decision/ homonyms/ metaphors), input modality (verbal/non-verbal, visual/auditory).

- ToM: type of ToM inference (cognitive/affective), stimulus type (animations/ cartoon/ photos/ RMET/ video/ stories), task type (e.g., false belief reasoning, predicting behaviours based on mental states), input modality, difficulty based on accuracy scores (experimental condition harder/easier than or equally difficult to the control condition).

- Empathy: target of empathy (e.g., pain, happiness, sadness etc.), valence of target of empathy (positive/negative), input modality, difficulty based on accuracy scores.

- Trait inference: type of inference (prior person knowledge/appearance/ behaviour), trait (e.g., trustworthiness, approachability etc.), input modality, difficulty based on accuracy scores.

- Moral reasoning: input modality, difficulty based on accuracy scores.

S.1.4. Missing information

If activation coordinates from whole-brain analyses for the contrast of interest were not reported in the published article or in supplementary materials, the authors of the study were contacted to obtain the missing information. If this was unsuccessful, the study was excluded from the meta-analysis. If the data was obtained directly from personal communication with the authors, this is clearly indicated in section 2 of the supplementary materials.

S.1.5. Overview of the study selection process

In accordance with the PRISMA guidelines (Page et al., 2021), a detailed overview of the study selection process is provided below separately for each dataset. Figures S1- S5 depict the number of articles identified and screened at each stage, the number of articles excluded at each stage, as well as the final number of articles and experiments included in the analyses.
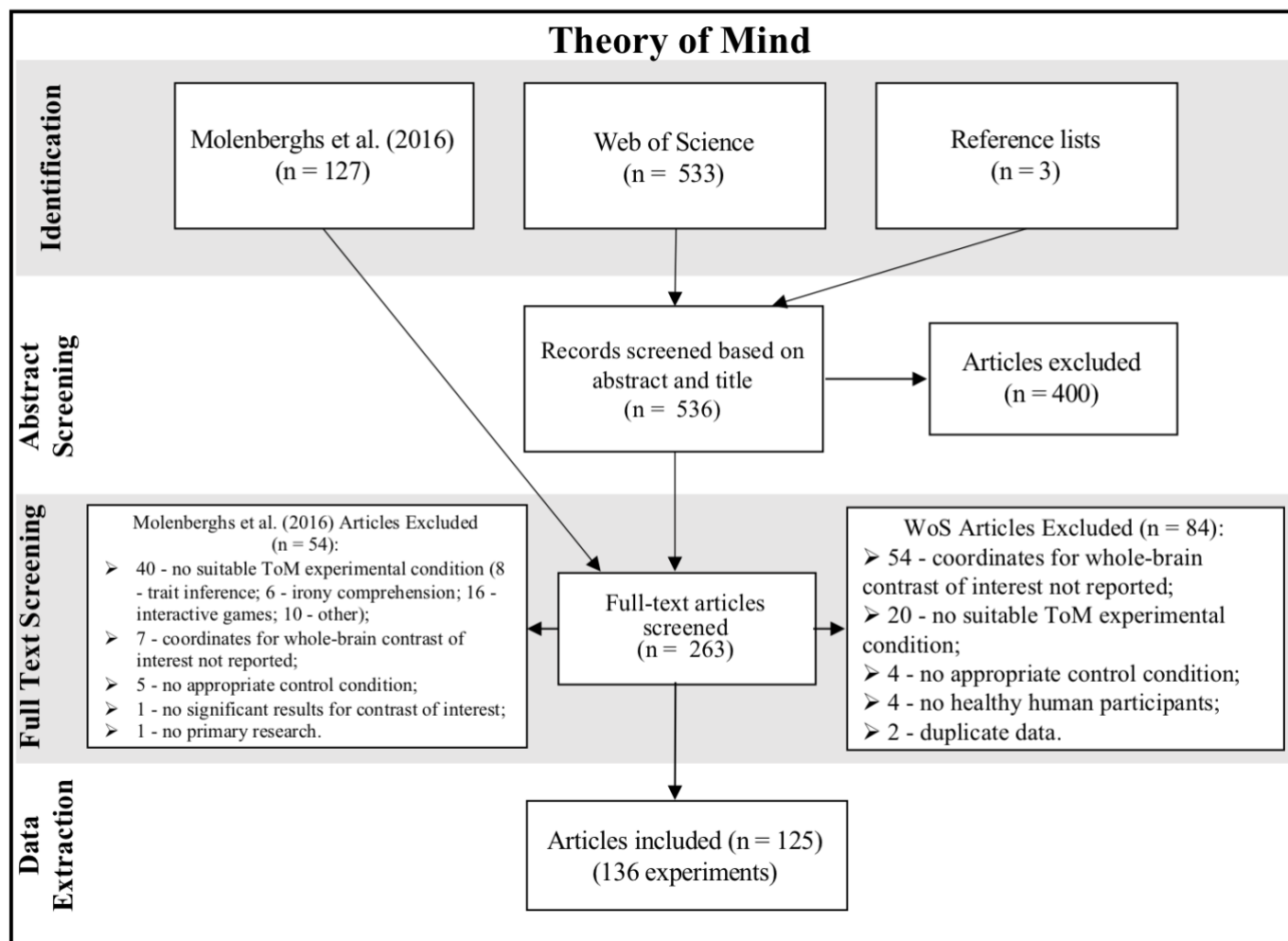
Figure S1. Overview of the selection process for the **theory of mind** meta-analysis. We included 136 experiments with a total number of 2158 coordinates and 3452 participants.
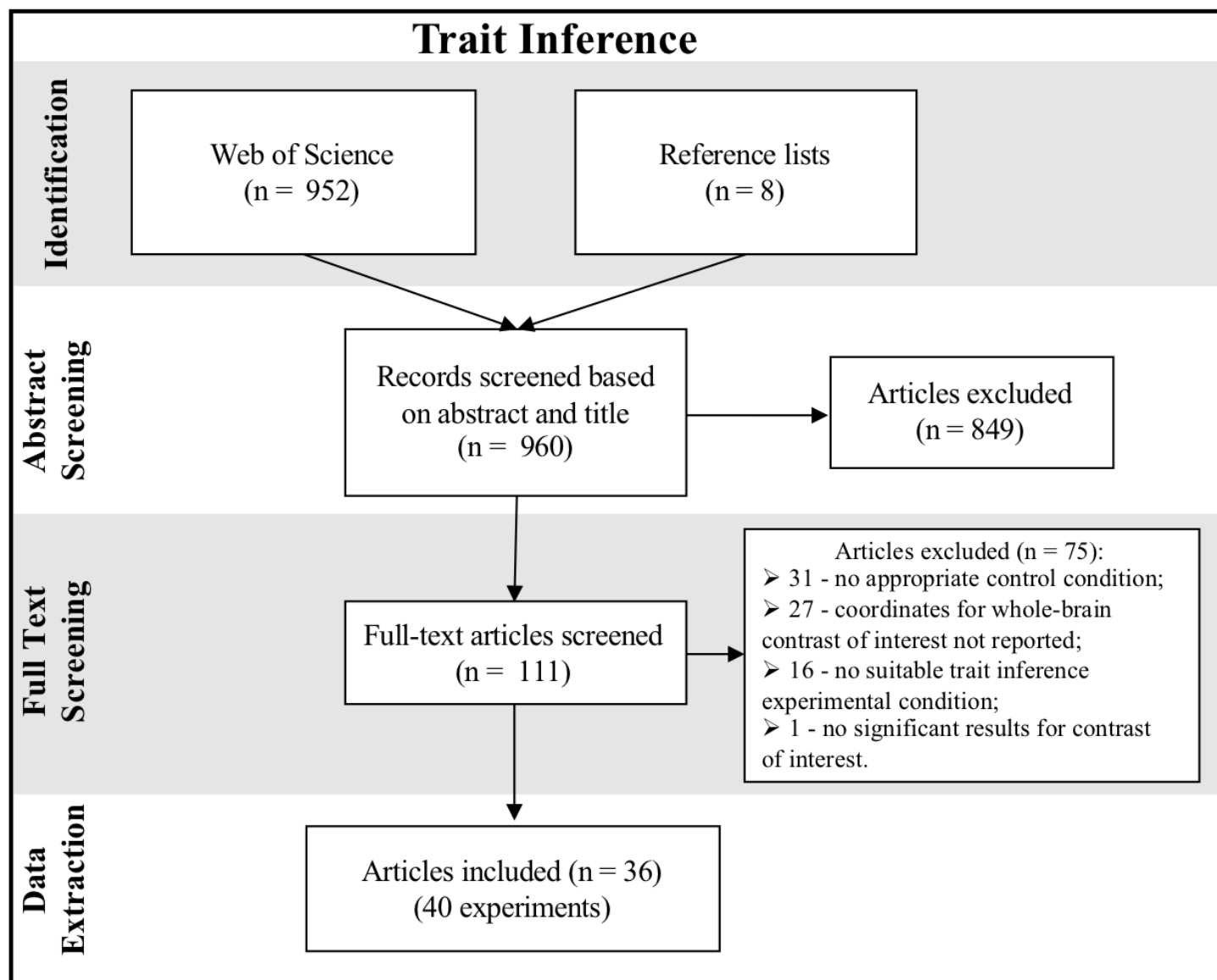
Figure S2. Overview of the selection process for the **trait inference** meta-analysis. We included 40 experiments with a total number of 523 coordinates and 732 participants.
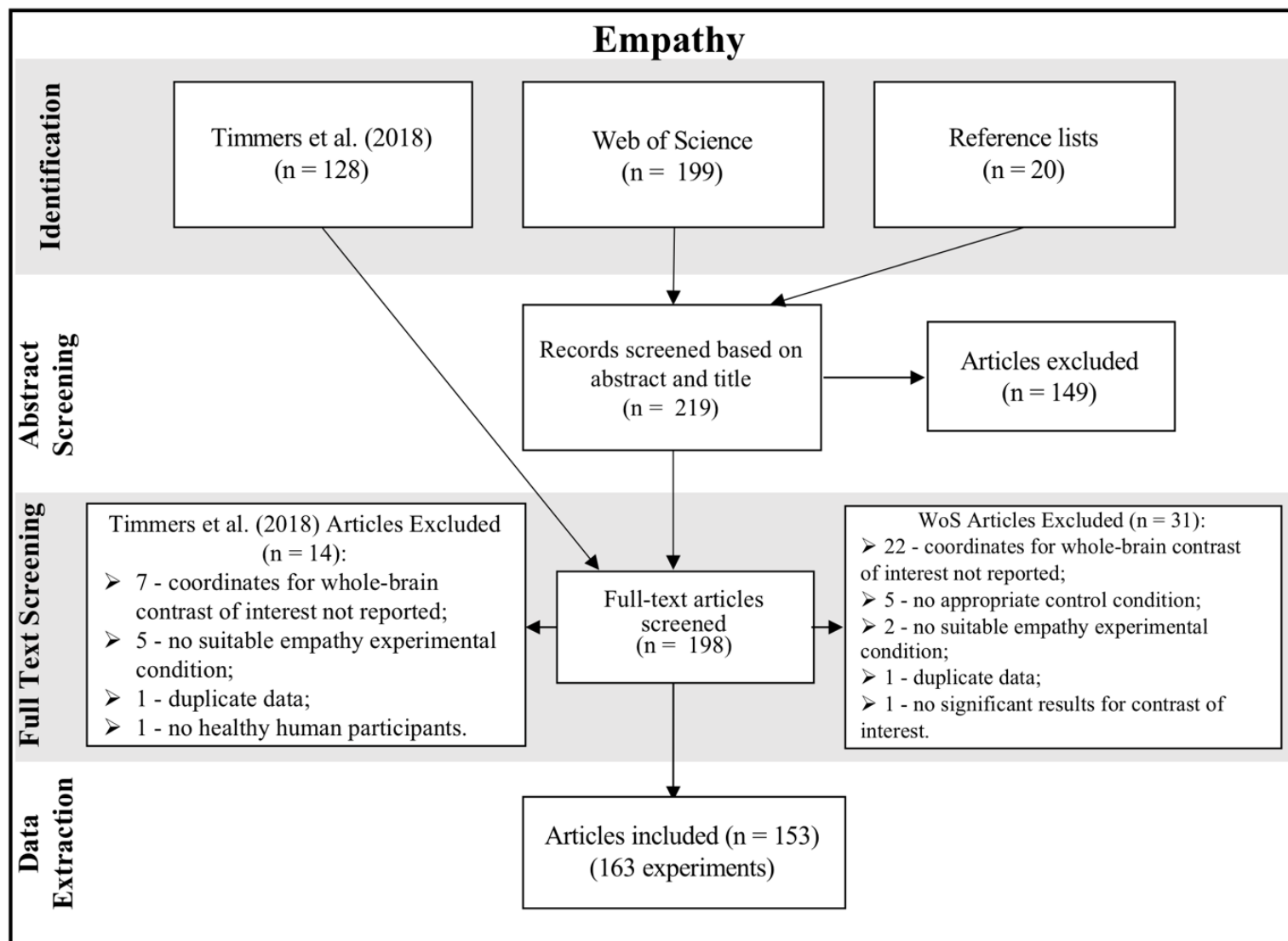
Figure S3. Overview of the selection process for the **empathy** meta-analysis. We included 163 experiments with a total number of 2691 coordinates and 4406 participants.
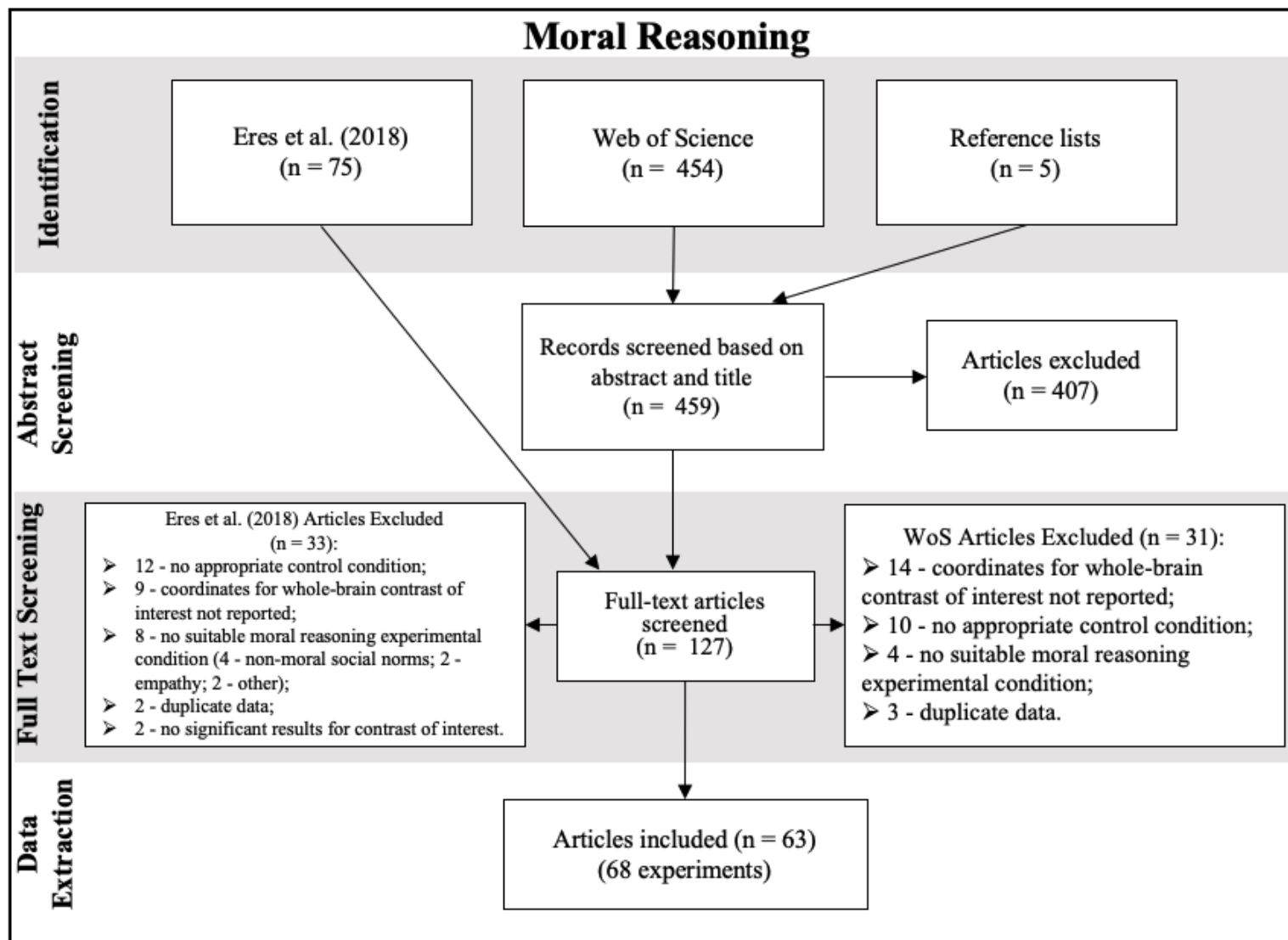
Figure S4. Overview of the selection process for the **moral reasoning** meta-analysis. We included 68 experiments with a total number of 884 coordinates and 1587 participants.
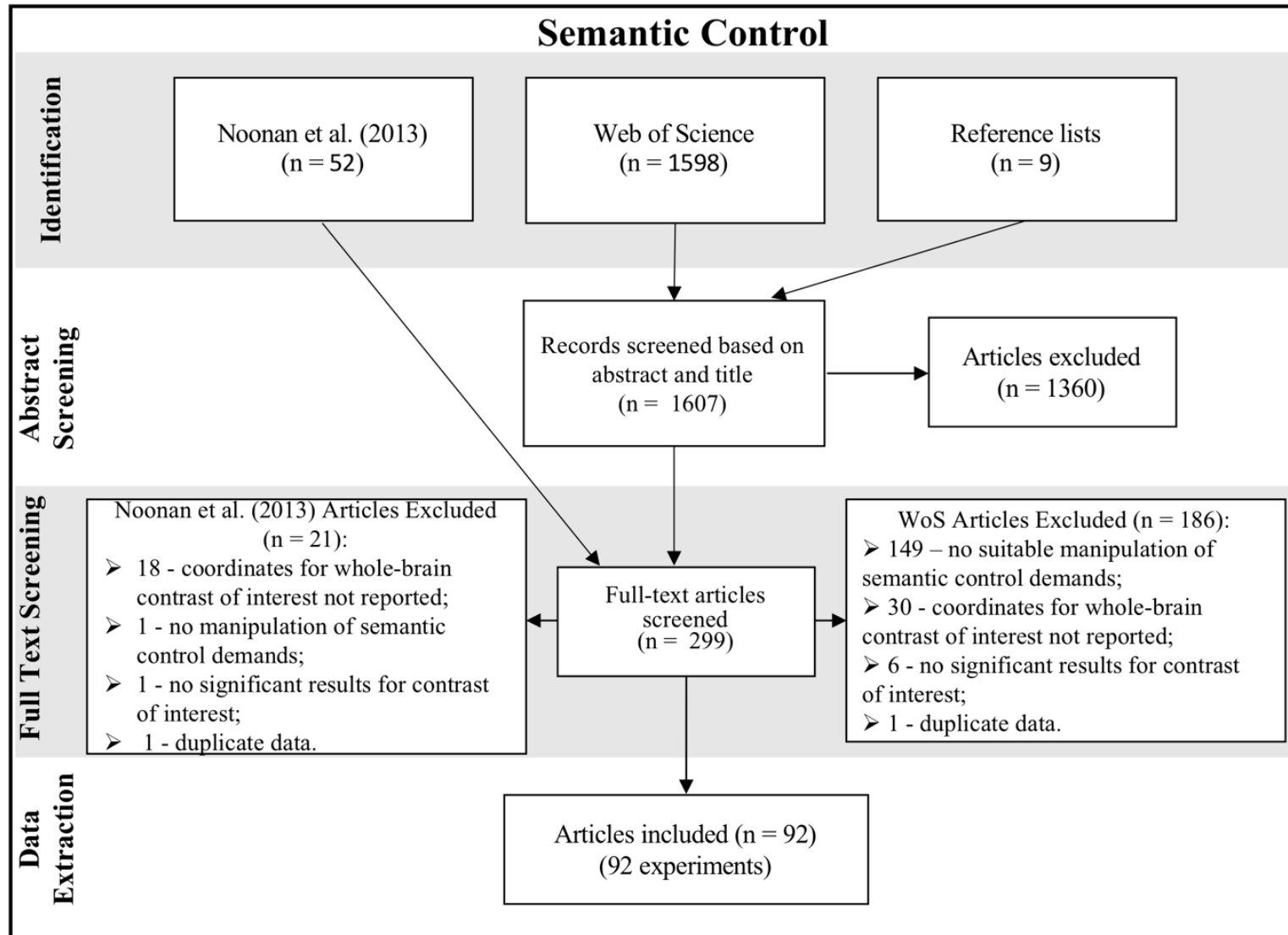
Figure S5. Overview of the selection process for the **semantic control** meta-analysis. We included 92 experiments with a total number of 971 coordinates and 1966 participants.

Section S2. Data analysis

A list of all analyses conducted is provided in the table below. For each analysis, the table specifies the type of analysis (C - confirmatory/ E - exploratory), the results section where the findings are described, the results table number and the results figure number.

Table S2. List of all analyses conducted.

Analysis			Type	Results Section	Results Table	Figure
Semantic control			C	3.2	S2	3
Social cognition	Independent sub-domains	Low baselines excluded	C	OSF page		-
		Theory of mind — Global Theory of Mind (all tasks) *	C	3.1.1	S1.1.1	1a
		Theory of mind — False belief reasoning*	E	S3.1	S1.1.2	S6a
		Trait Inference*	C	3.1.2	S1.2	1b
		Empathy — Global Empathy*	C	3.1.3	S1.3.1	S7a
		Empathy — Empathy for Pain*	E	3.1.3	S1.3.2	1d
		Empathy for Emotions*	E	3.1.3	S1.3.3	1c
		Empathy for Pain vs. Emotions^	E	3.1.3	S1.3.4	S7b
		Moral reasoning*	C	3.1.4	S1.4	1e
	Overlay conjunction		E	S3.3	2.1.5	2; S7
	Explicit vs. Implicit	Empathy for Emotions — Explicit*	C	S3.4	S4.1.1	5a
		Empathy for Emotions — Implicit*	C	S3.4	S4.1.2	5a
		Empathy for Emotions — Explicit vs. Implicit^	C	S3.4	S4.2.1	5d
		Empathy for Pain — Explicit*	C	S3.4	S4.1.3	5b
		Empathy for Pain — Implicit*	C	S3.4	S4.1.4	5b
		Empathy for Pain — Explicit vs. Implicit^	C	S3.4	S4.2.2	5e
		Moral reasoning — Explicit*	C	S3.4	S4.1.5	5c
		Moral reasoning — Implicit*	C	S3.4	S4.1.6	5c
		Moral reasoning — Explicit vs. Implicit^	C	S3.4	S4.2.3	5f
		Cluster analyses	E	3.4.2	-	S8
	Task difficulty	E>C ToM*	E	3.5	S5.1	6a
		E=C ToM*	E	3.5	S5.2	6a
		C>E ToM*	E	3.5	S5.4	6a
		E>C vs. E=C^	E	3.5	S5.3	-
		ToM known difficulty category*	E	-	S5.5	6c
		Cluster analysis	E	3.5	-	6c
Social	Theory of Mind vs. SC	Global ToM (all tasks) vs. SC^	C	3.3.1	S3.1	4a
		False belief reasoning vs. SC^	E	S3.1	S3.1.2	S6b

	Trait inference vs. SC^		C	3.3.2	S3.2	4b
	Empathy vs. SC	Global Empathy vs. SC^	C	OSF project		
		Empathy for Emotions vs. SC^	E	3.3.3	S3.3	4c
		Empathy for Pain vs. SC^	E	3.3.4	S3.4	4d
	Moral reasoning vs. SC^		C	3.3.5	S3.5	4e

* Indicates independent ALE analyses. Each independent ALE analysis was repeated using two different statistical thresholding approaches (see the Results section).

^ Indicates contrast and conjunction analyses.

SC = Semantic control; E = exploratory analysis; C = confirmatory analysis

Section S3. Results

S3.1. False belief reasoning

The validity of some of the tasks used to investigate ToM processing has been debated (Heyes, 2014; Oakley et al., 2016; Obhi, 2012; Quesque and Rossetti, 2020). For example, Quesque & Rossetti (2020) argue that some classic ToM tasks (e.g., Heider and Simmel) do not necessarily require the representation of other's mental states and successful completion can be explain by lower-level cognitive processes (e.g., basic associative learning mechanisms). A further criticism has been that they do not necessitate distinguishing one's own mental states from others' mental states (Heyes, 2014). False belief tasks, on the other hand, are frequently used in both healthy and clinical populations (e.g., Yirmiya et al., 1998), as well as in developmental contexts (e.g., Wellman et al., 2001), and are more commonly accepted as a suitable task for probing ToM and for identifying the underlying brain network (Dodell-Feder et al., 2011; Saxe and Kanwisher, 2003). False belief tasks require participants to make inferences about the (false) beliefs of protagonists which may contrast their own knowledge about the state of reality, meeting both aforementioned criteria for valid ToM tasks (Quesque and Rossetti, 2020).

Our global ToM meta-analysis pooled across a variety of ToM task paradigms, not all of which may be considered valid measures of ToM processing. Therefore, we performed a separate analysis focused solely on experiments which employed false belief tasks to explore any differences in the underlying brain network when using a more conservative task definition. Across 43 false belief reasoning experiments, the ALE analysis revealed convergent activation in 9 clusters including in the precuneus, bilateral pMTG & STG, bilateral mid MTG

(extending anteriorly towards the temporal pole), right MFG, mPFC, medial OFC and cerebellum (Figure S6a; Table S1.1.2). All these clusters survived both extent-based and height-based thresholding. Unlike in the inclusive ToM analysis, we did not find convergent activation in the bilateral IFG. Overlap between the neural network underpinning semantic control (i.e., SCN & regions of the MDN) and false belief reasoning was found in a single left pMTG cluster (Figure S6b; Table S3.1.2).
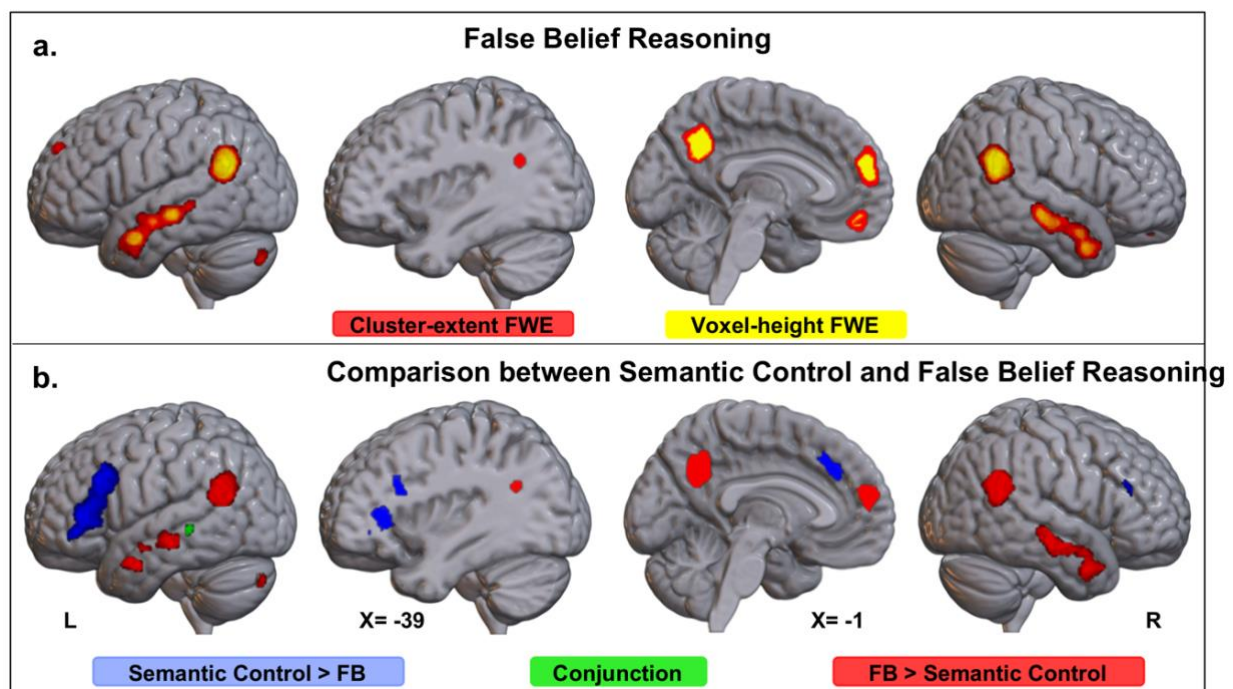

Figure S6. a) Binary whole-brain ALE maps showing statistically significant convergent activation resulting from independent meta-analyses of false belief reasoning experiments (N=43). The ALE maps were thresholded using an FWE corrected cluster-extent at $p < .05$ with a cluster-forming threshold of $p < .001$ (red) and, an FWE corrected voxel-height threshold of $p < .05$ (yellow). b) Results of the contrast (blue: semantic control > false belief reasoning; red: false belief reasoning > semantic control) and conjunction (green) analyses between the ALE maps associated with false belief reasoning and semantic control. The contrast maps were thresholded with a cluster-forming threshold of $p < .001$ and a minimum cluster size of 200 mm^3. The lateral views, which show projections on the cortical surface, are accompanied by brain slices at the sagittal midline and also coplanar with the peak of the left IFG [X = -39] cluster that overlapped across all social domains (Table S1.5). FB = false belief reasoning.
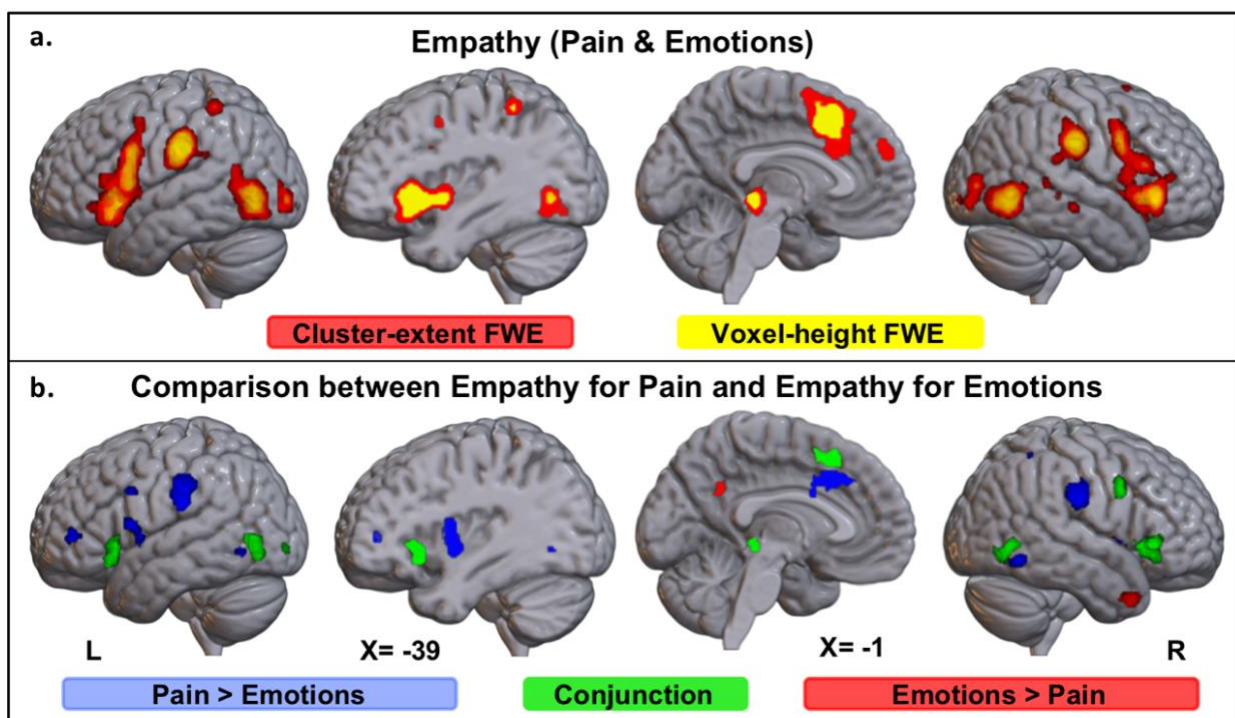
S3.2. Global Empathy (for Pain & Emotions)

Figure S7. a) Binary whole-brain ALE maps showing statistically significant convergent activation resulting from independent meta-analyses of all empathy studies (N=163), including both empathy for pain and empathy for emotions. The ALE maps were thresholded using an FWE corrected cluster-extent at $p < .05$ with a cluster-forming threshold of $p < .001$ (red) and, an FWE corrected voxel-height threshold of $p < .05$ (yellow). b) Results of the contrast (blue: empathy for pain > empathy for emotions; red: empathy for emotions > empathy for pain) and conjunction (green) analyses between the ALE maps associated with empathy for pain and empathy for emotions. The contrast maps were thresholded with a cluster-forming threshold of $p < .001$ and a minimum cluster size of 200 mm^3. The lateral views, which show projections on the cortical surface, are accompanied by brain slices at the sagittal midline and also coplanar with the peak of the left IFG cluster that overlapped across all social domains [Table S1.5; X = -39]).

S3.3. A common network for multiple sub-domains of social cognition

To identify brain areas consistently activated across multiple sub-domains of social cognition, we performed an overlay conjunction analysis of the ALE maps associated with ToM, trait inference, empathy (for pain and/or emotions) and moral reasoning. The ALE maps were thresholded using p<.001 uncorrected. Convergent activation across all four socio-cognitive sub-domains was found in the bilateral IFG (pars orbitalis), mPFC, precuneus, left pSTG, and bilateral ATL (Figure S8).
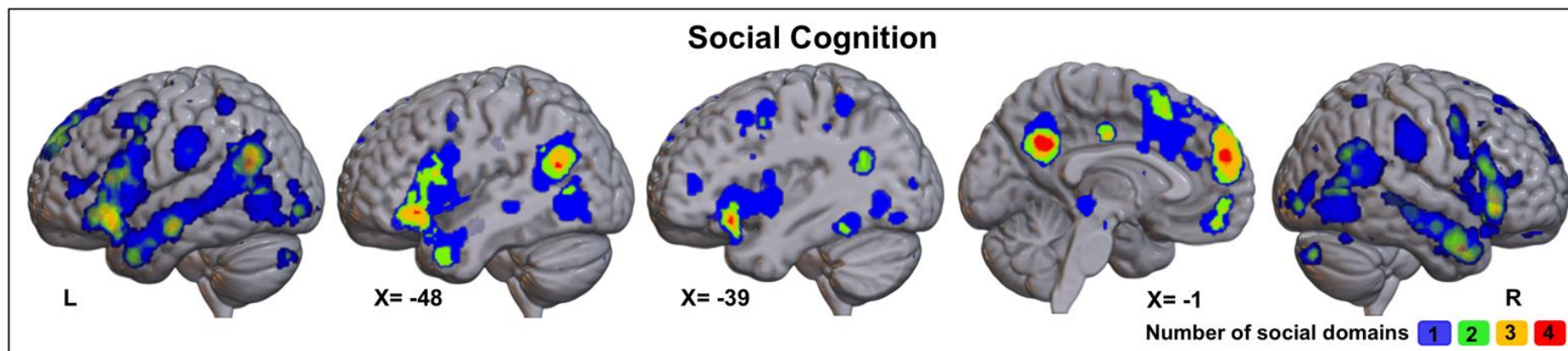

Figure S8. An overlay conjunction of the ALE maps resulting from independent meta-analyses on ToM, trait inference, empathy for pain/emotions, and moral reasoning. The map displays the number of social domains showing convergent activation in each voxel. The ALE maps were thresholded using an uncorrected cluster-forming threshold of $p < .oo1$. The lateral views, which show projections on the cortical surface, are accompanied by brain slices at the sagittal midline and also coplanar with the peak of the left pSTG (X = -48) and left IFG (X = -39) clusters that overlapped across all four social domains.

S3.4. Explicit versus Implicit Social Paradigms

S3.4.1. Conjunctions and Contrasts between Explicit and Implicit Experiments

<u>Empathy for affective states</u>

An ALE analysis across all explicit paradigms showed convergent activation in the bilateral IFG, mPFC, SMA, precuneus, middle cingulate gyrus, bilateral temporal pole, left caudate and brainstem. An ALE analysis of all implicit experiments revealed convergent activation in the bilateral IFG, bilateral pMTG and right occipital cortex. The conjunction analysis between implicit and explicit empathy for affective states revealed overlapping activation in two clusters located in pars triangularis of the right IFG and pars orbitalis of the left IFG. Formal contrasts did not reveal any above-threshold differential activation.

<u>Empathy for pain</u>

The ALE analysis of empathy for pain studies employing explicit instructional cues revealed convergent activation in the bilateral insula (extending to the IFG), pars triangularis of left IFG (extending to MFG), the middle cingulate gyri (extending to SMA), bilateral SMG, left inferior occipital gyrus (extending to pMTG) left thalamus and right pallidum. An ALE analysis of studies using an implicit paradigm revealed convergent activation in the bilateral IFG, bilateral precentral gyrus, bilateral insula, SMA, middle cingulate gyrus, bilateral SMG, left IPL, bilateral posterior ITG, calcarine fissure, left thalamus, right amygdala and brainstem. The conjunction between implicit and explicit empathy for pain revealed overlapping activation in the left IFG (pars opercularis), bilateral insula, dmPFC, middle cingulate cortex, SMA, bilateral SMG and left inferior occipital gyrus. Formal contrasts found that explicit paradigms differentially engage dmPFC while implicit paradigms show additional convergent activation in left inferior occipital cortex and right hippocampus.

<u>Moral reasoning</u>

Convergent activation across explicit paradigms was found in the left insula cortex, medial OFC, dmPFC, precuneus and left pMTG. Convergent activation across implicit paradigms was found in the left IFG (pars opercularis), dmPFC and right anterior MTG. The conjunction between implicit and explicit moral reasoning paradigms revealed overlapping activation in mPFC, while the formal contrasts did not reveal any above-threshold differential activation.

3.4.2. Cluster analyses investigating the contribution of explicit/implicit experiments to ALE clusters

In addition, we conducted exploratory cluster analyses to investigate whether the explicit and implicit experiments contributed similarly to each of the significant ALE clusters found for each social domain. Given that the sample sizes of implicit experiments were too small to conduct independent ALE analyses for the ToM and Trait Inference datasets, this cluster analysis allowed us to understand whether a specific type of instructional cue drove the convergent activation in any of the identified clusters. For each social domain, we calculated the proportion of implicit and explicit experiments that contributed at least one peak to each identified cluster. The results (Figure S9) revealed that explicit and implicit paradigms contributed equally to most activation clusters (all $p > .05$), with the exception of the precuneus cluster in response to ToM ($p = .023$, explicit > implicit) and a cluster located in the right primary visual cortex in response to empathy for pain ($p = .016$, implicit > explicit). This suggests that implicit and explicit experiments contributed equally to most activation clusters regardless of social domain.
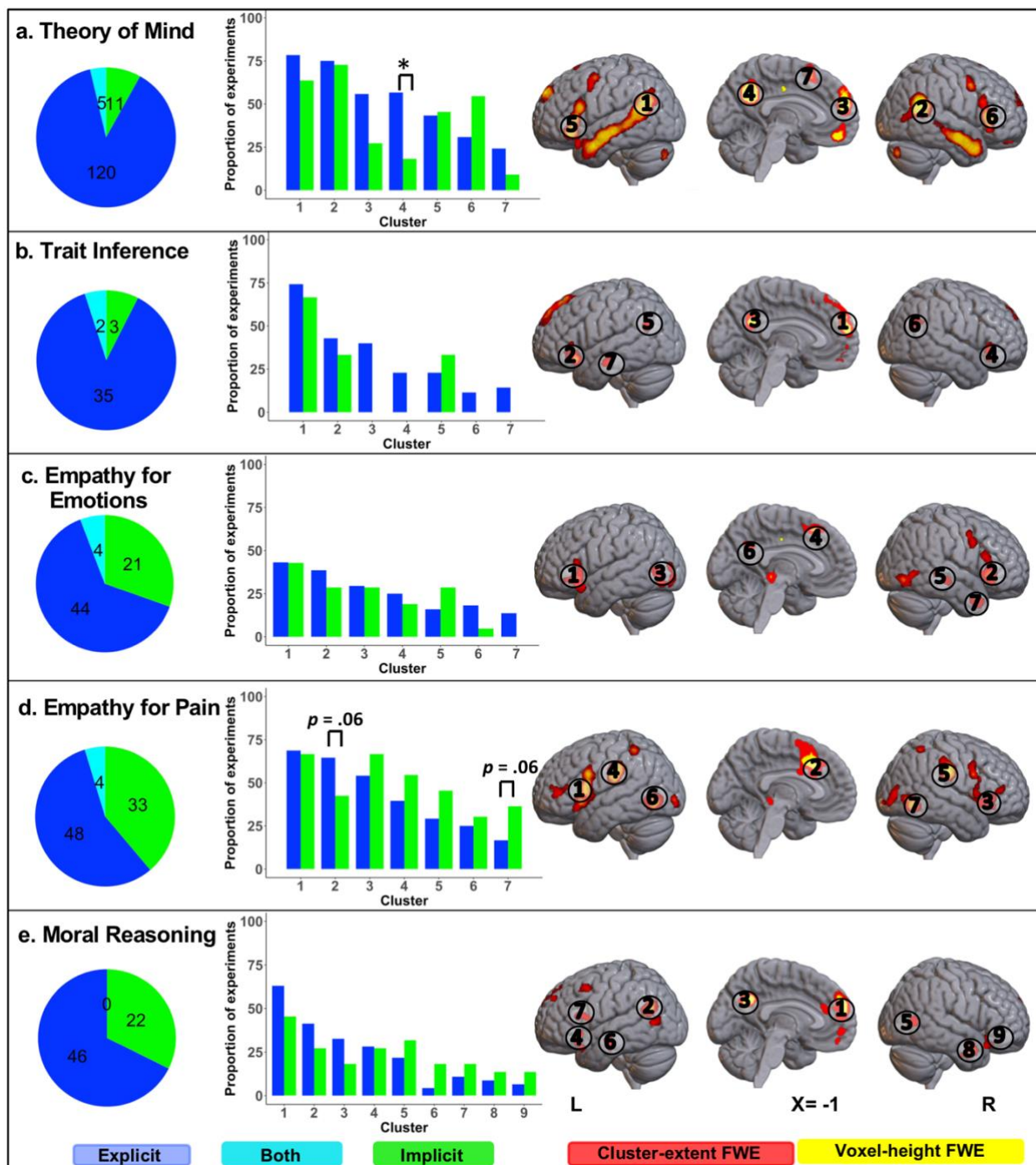
Figure S9. Pie charts illustrate the proportion of explicit, implicit and both (i.e., contrasts reporting coordinates from both implicit and implicit instructions) experiments that were included in the meta-analysis for each social domain. The labels indicate the experiment count for each category. The bar charts display the proportion of experiments in each category (explicit/implicit) that contributed at least one peak to the clusters of interest. The ALE map of convergent activation for each social domain is displayed on a standard MNI brain and each cluster of interest is indicated and numbered. * $p < .05$

S3.5. The difficulty characteristics of the explicit experiments included in the dataset for each social domain
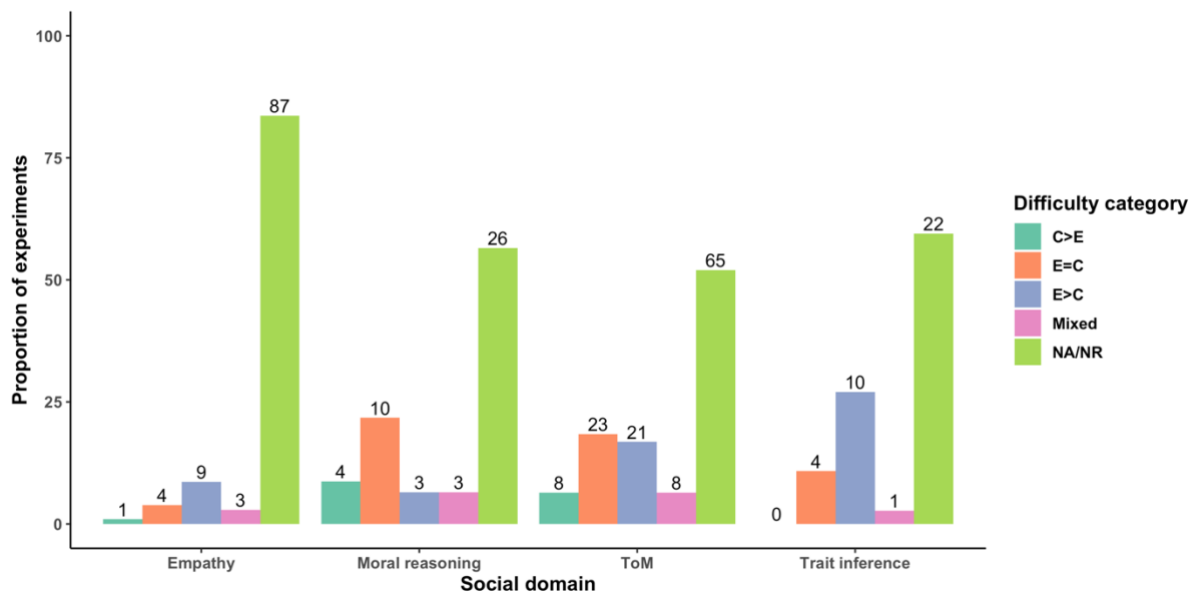

Figure S10. For each social domain, the proportion of studies featuring specific difficulty characteristics is illustrated. The bars represent the proportion of studies featuring specific difficulty characteristics. The exact number of experiments featuring specific difficulty characteristics is also displayed for each bar. C>E – reaction times were significantly faster in the experimental condition compared to the control condition suggesting that the control condition was more difficult. E=C – the was no statistically significant different between the reaction times in the experimental and control conditions suggesting equally difficult conditions. E>C – reaction times were significantly faster in control condition compared to the experimental condition suggesting that the experimental condition was more difficult. Mixed – the experiment pooled across contrasts belonging to different difficulty categories. NA – not applicable, NR – not reported.

References

Ding, R., Ren, J., Li, S., Zhu, X., Zhang, K., Luo, W., 2020. Domain-general and domain-preferential neural correlates underlying empathy towards physical pain, emotional situation and emotional faces: An ALE meta-analysis. Neuropsychologia 137, 107286. https://doi.org/10.1016/j.neuropsychologia.2019.107286

Dodell-Feder, D., Koster-Hale, J., Bedny, M., Saxe, R., 2011. FMRI item analysis in a theory of mind task. Neuroimage 55, 705–712. https://doi.org/10.1016/j.neuroimage.2010.12.040

Eickhoff, S.B., Nichols, T.E., Laird, A.R., Hoffstaedter, F., Amunts, K., Fox, P.T., Bzdok, D., Eickhoff, C.R., 2016. Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. Neuroimage 137, 70–85. https://doi.org/10.1016/j.neuroimage.2016.04.072

Eres, R., Louis, W.R., Molenberghs, P., 2018. Common and distinct neural networks involved in fMRI studies investigating morality: an ALE meta-analysis. Soc. Neurosci. 13, 384–398. https://doi.org/10.1080/17470919.2017.1357657

Heyes, C., 2014. Submentalizing. Perspect. Psychol. Sci. 9, 131–143. https://doi.org/10.1177/1745691613518076

Kogler, L., Müller, V.I., Werminghausen, E., Eickhoff, S.B., Derntl, B., 2020. Do I feel or do I know? Neuroimaging meta-analyses on the multiple facets of empathy. Cortex 129, 341–355. https://doi.org/10.1016/j.cortex.2020.04.031

Molenberghs, P., Johnson, H., Henry, J.D., Mattingley, J.B., 2016. Understanding the minds of others: A neuroimaging meta-analysis. Neurosci. Biobehav. Rev. 65, 276–291. https://doi.org/10.1016/j.neubiorev.2016.03.020

Oakley, B.F.M., Brewer, R., Bird, G., Catmur, C., 2016. Theory of mind is not theory of emotion: A cautionary note on the reading the mind in the eyes test. J. Abnorm. Psychol. 125, 818–823. https://doi.org/10.1037/abn0000182

Obhi, S.S., 2012. The Amazing Capacity to Read Intentions from Movement Kinematics. Front. Hum. Neurosci. 6, 162. https://doi.org/10.3389/fnhum.2012.00162

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A.,

Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372, n71. https://doi.org/10.1136/bmj.n71

Quesque, F., Rossetti, Y., 2020. What Do Theory-of-Mind Tasks Actually Measure? Theory and Practice. Perspect. Psychol. Sci. 15, 384–396. https://doi.org/10.1177/1745691619896607

Saxe, R., Kanwisher, N., 2003. People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind." Neuroimage 19, 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Timmers, I., Park, A.L., Fischer, M.D., Kronman, C.A., Heathcote, L.C., Hernandez, J.M., Simons, L.E., 2018. Is Empathy for Pain Unique in Its Neural Correlates? A Meta-Analysis of Neuroimaging Studies of Empathy. Front. Behav. Neurosci. 12, 289. https://doi.org/10.3389/fnbeh.2018.00289

Wellman, H.M., Cross, D., Watson, J., 2001. Meta-analysis of theory-of-mind development: The truth about false belief. Child Dev. 72, 655–684. https://doi.org/10.1111/1467-8624.00304

Yirmiya, N., Erel, O., Shaked, M., Solomonica-Levi, D., 1998. Meta-Analyses Comparing Theory of Mind Abilities of Individuals with Autism, Individuals with Mental Retardation, and Normally Developing Individuals. Psychol. Bull. 124, 283–307. https://doi.org/10.1037/0033-2909.124.3.283

APPENDIX III. Supplementary Materials

for

CHAPTER 5. Graded functional organisation in the left inferior frontal gyrus: evidence from task-free and task-based functional connectivity
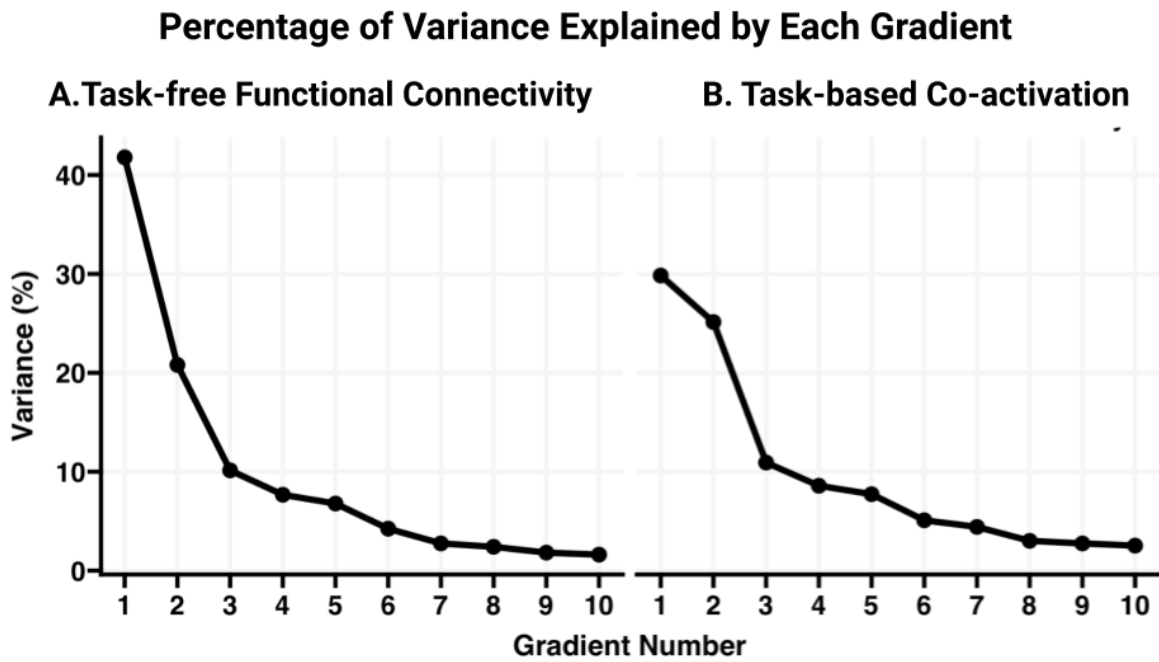

Figure S1. Percentage of variance explained by the 10 gradients derived from A) the task-free functional connectivity data and B) the task-based co-activation patterns.
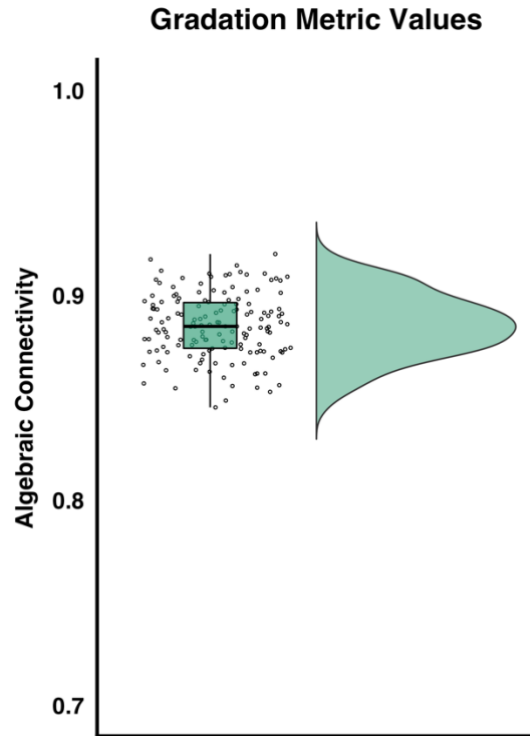
Gradation Metric Values

Figure S2. The distribution of the algebraic connectivity values (equivalent to the second largest eigenvalue of the Laplacian of the similarity matrix) obtained per participant in the task-free functional connectivity assessment are illustrated alongside individual datapoints and a boxplot highlighting the median, 25th and 75th quartiles. Values near 0 reflect the existence of hard clusters, whereas higher numbers suggest a graded change in functional connectivity. Note that the y-axis starts at 0.7, which is above the midpoint of possible values. The individual-level gradation metric values suggest that the left IFG is characterized by graded changes in task-free functional connectivity.
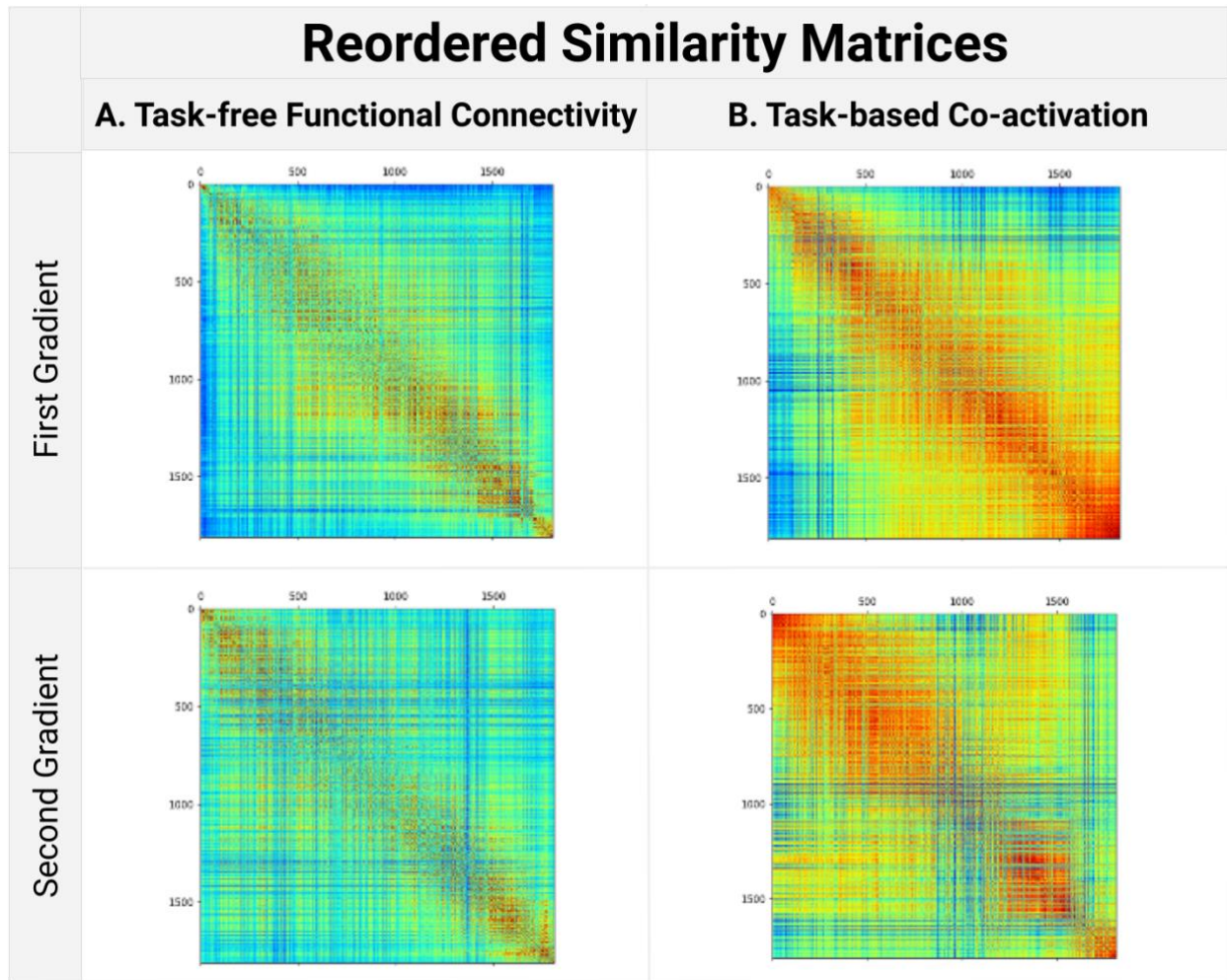
Figure S3. Similarity matrices reordered based on the voxels' positions along the first and second gradients. A) Reordered task-free FC group matrix. B) Reordered task-based co-activation matrix. Visual inspection of the reordered matrices suggests a high degree of gradation in the main axes of functional connectivity change across the left IFG.

Consistency Between the Task-free and Task-based Gradient Maps

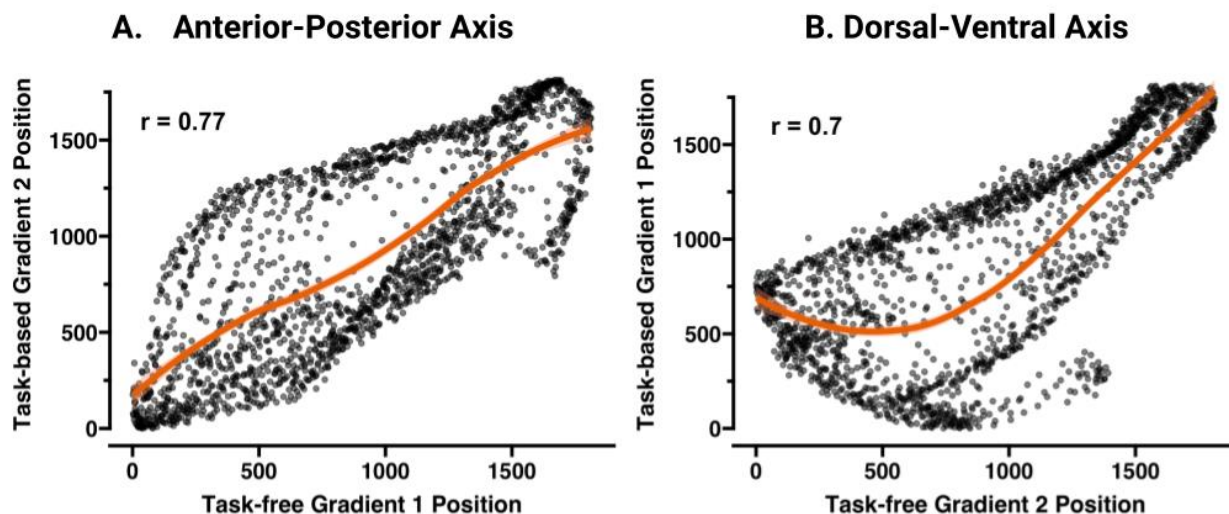**A. Anterior-Posterior Axis**

B. Dorsal-Ventral Axis

Figure S4. Scatterplots illustrate the relationship between voxels' positions in the resting-state gradients and the task-state gradients. The loess lines depicted in orange highlights the functional relationship. A. Voxels' ranks on the anterior-posterior (first) task-free gradient are plotted against their ranks on the anterior-posterior (second) task-based gradient. B. Voxels' ranks on the dorsal-ventral (second) task-free gradient are plotted against their ranks on the dorsal-ventral (first) task-based gradient. The r values represent the product-moment correlation coefficients and suggest strong relationships between the gradients extracted from independent FC datasets.
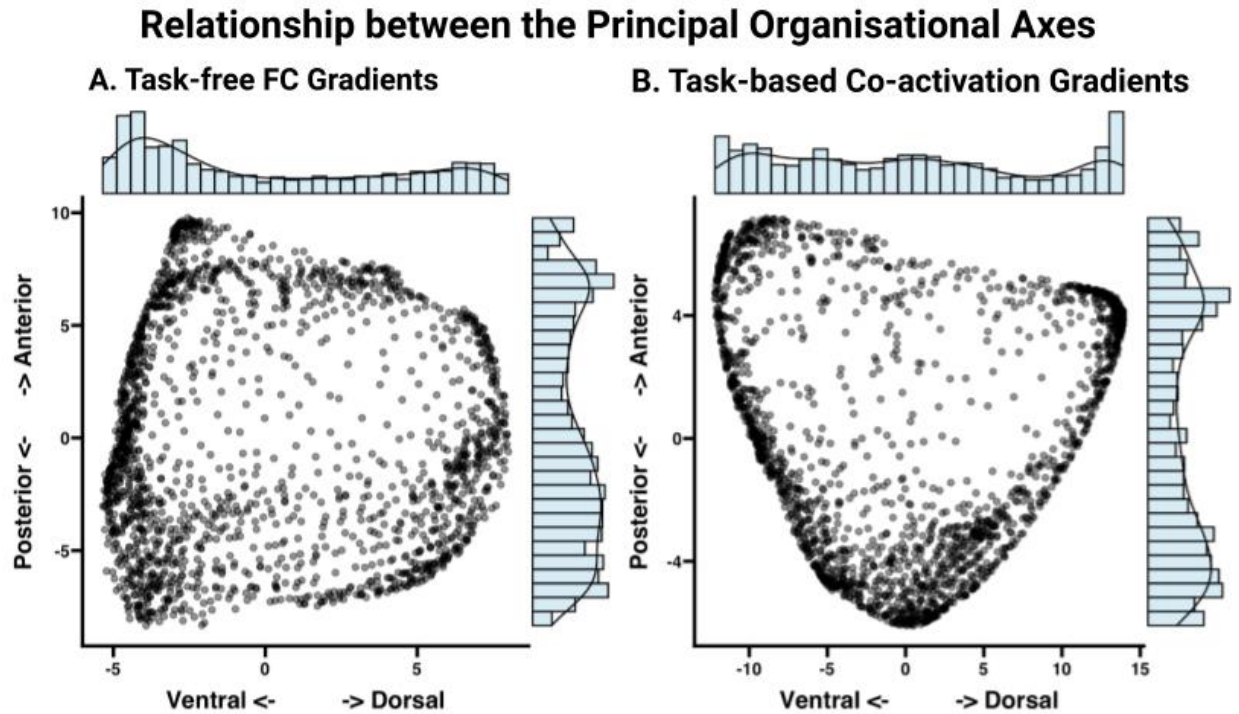
Figure S5. Scatterplots illustrate the relationship between voxels' gradient values on the first two connectivity embedding gradients extracted from A) task-free functional connectivity and B) task-based co-activation patterns. Histograms and density plots depicting the distribution of gradient values are presented on the respective axes.
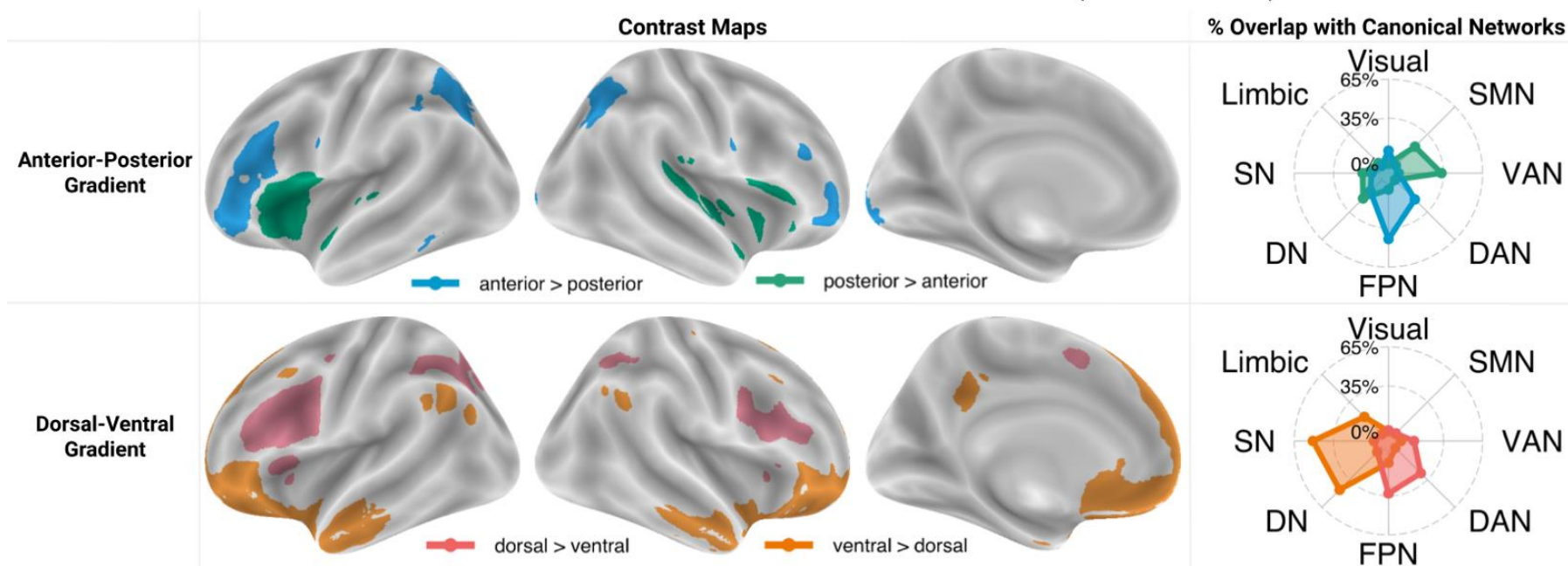
Figure S6. Results of contrast analyses between task-constrained co-activation patterns (derived using MACM analyses) of the IFG clusters located at the extremes of the anterior-posterior and dorsal-ventral task-based gradients. Unlike in the main text, these contrast maps were <u>not</u> masked using independent MACM maps. The spider plots in the right column show the percentage of overlap between the contrast maps and canonical networks from Yeo et al. (2011), as well as the semantic network from Jackson et al. (2016), which is comprised of regions that are functionally coupled with the ventrolateral anterior temporal lobe semantic hub at rest.
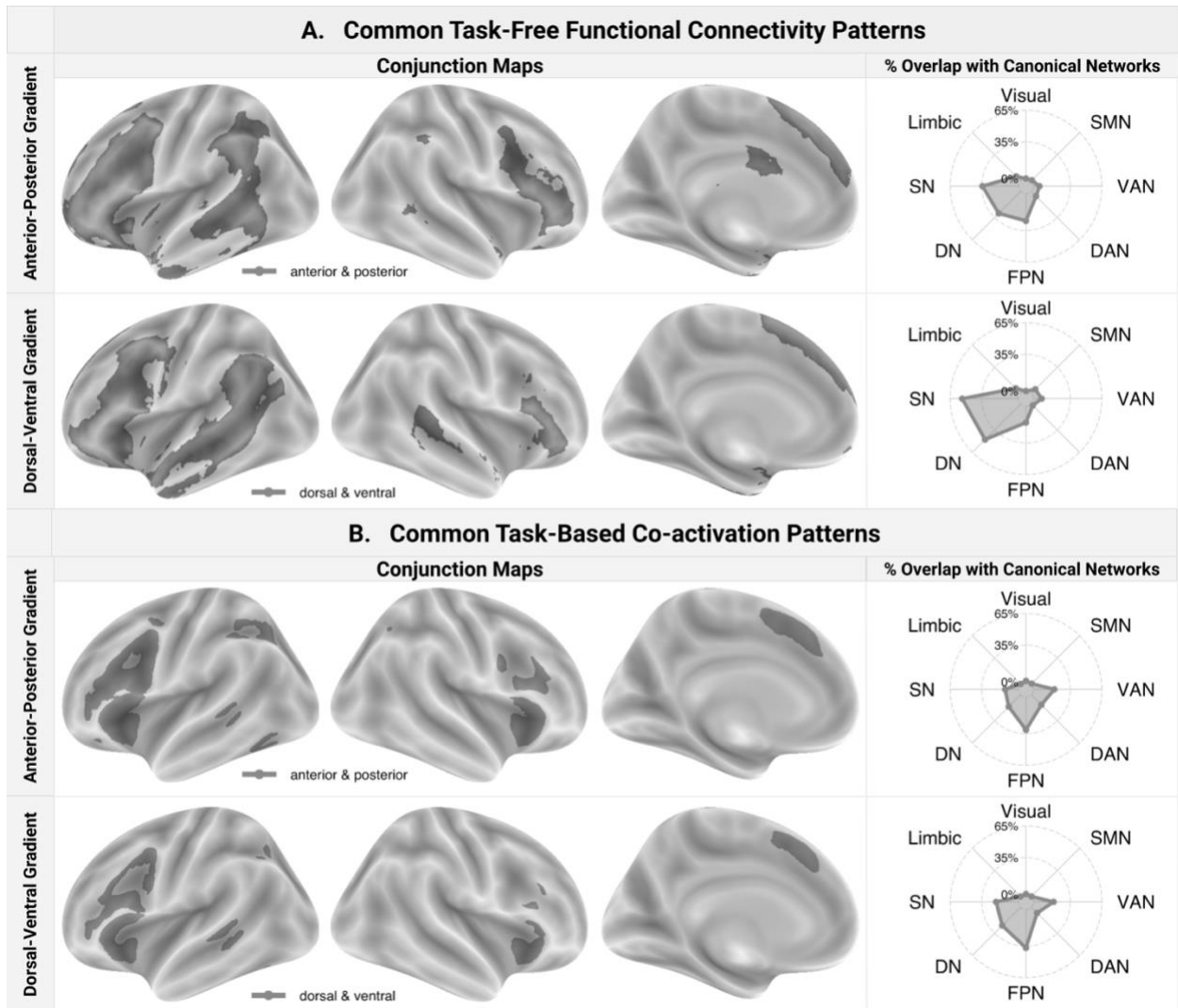
Figure S7. Conjunction maps showing common regions of (A) task-free functional connectivity and (B) co-activation between the hard clusters located at the extremes of the respective gradients. The spider plots in the right column show the percentage of overlap between the contrast maps and canonical networks.
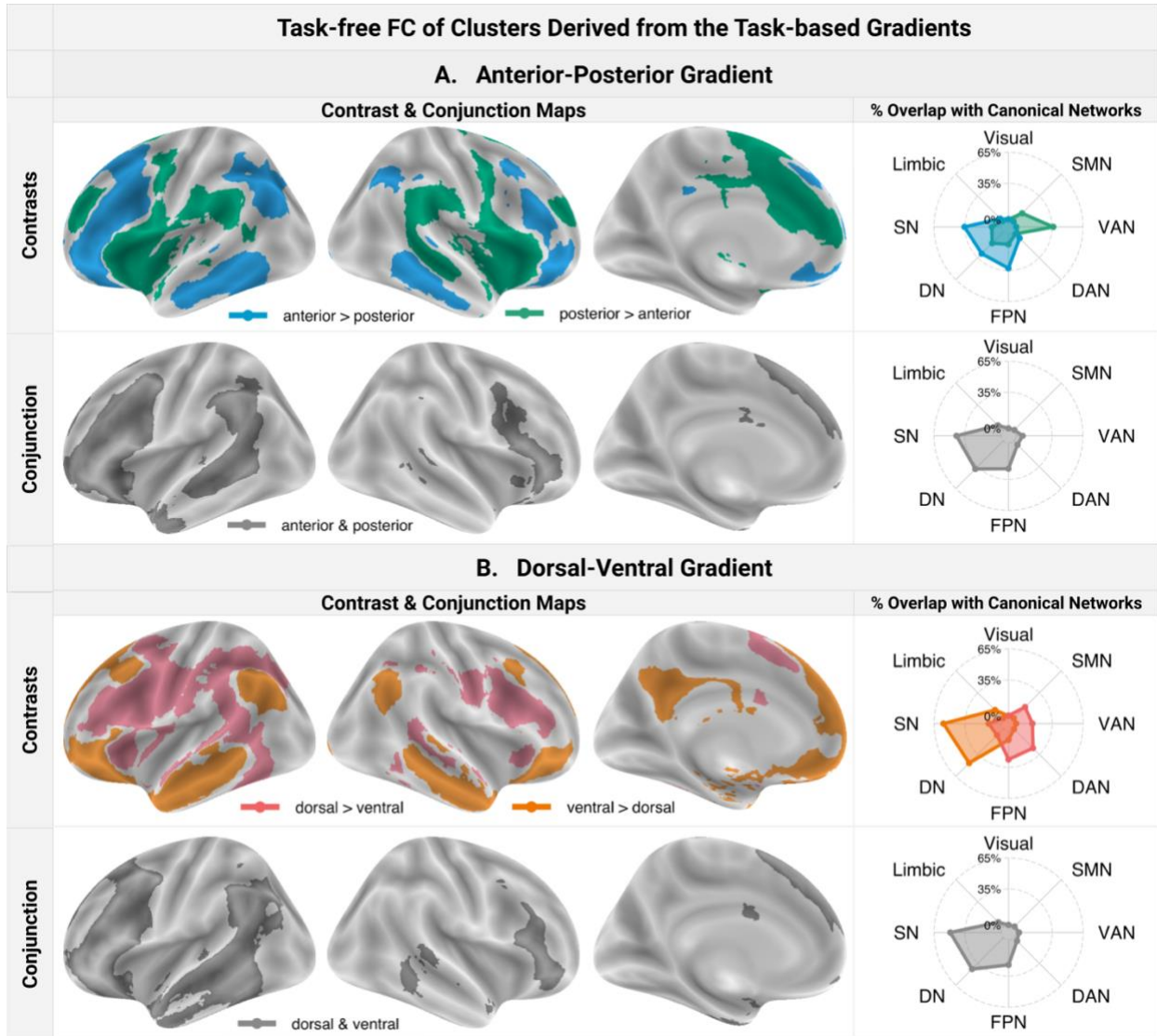
Figure S8. Contrast and conjunction maps showing regions of common and differential functionally coupling at rest between hard clusters representing the edges of the task-based FC (A) anterior-posterior gradient map and (B) dorsal-ventral gradient map. The spider plots in the right column show the percentage of overlap between the contrast maps and canonical networks.
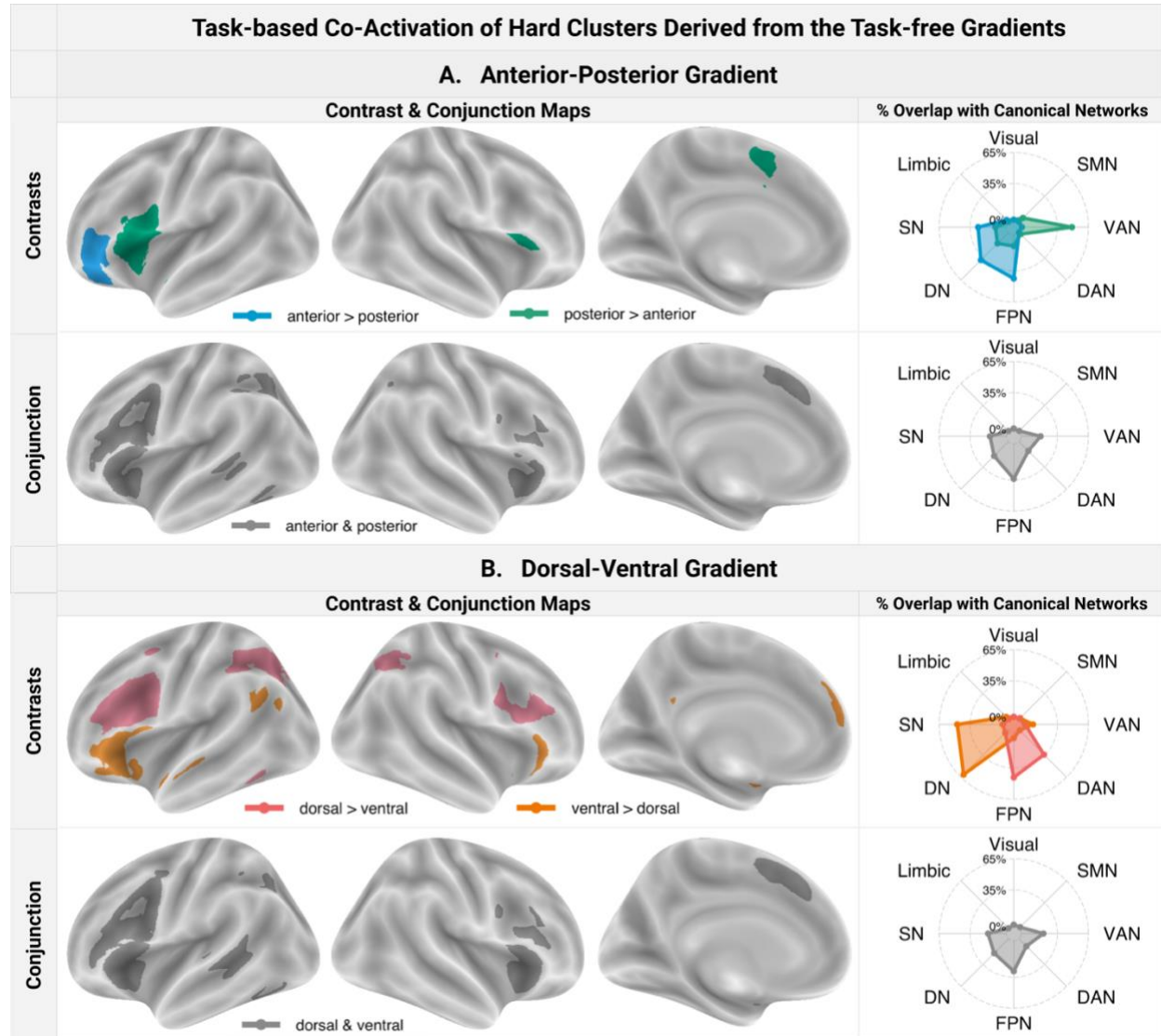
Figure S9. Contrast and conjunction maps showing regions of common and differential task-constrained co-activation across cognitive domains between hard clusters representing the edges of the task-free FC (A) anterior-posterior gradient map and (B) dorsal-ventral gradient map. The spider plots in the right column show the percentage of overlap between the contrast maps and canonical networks.
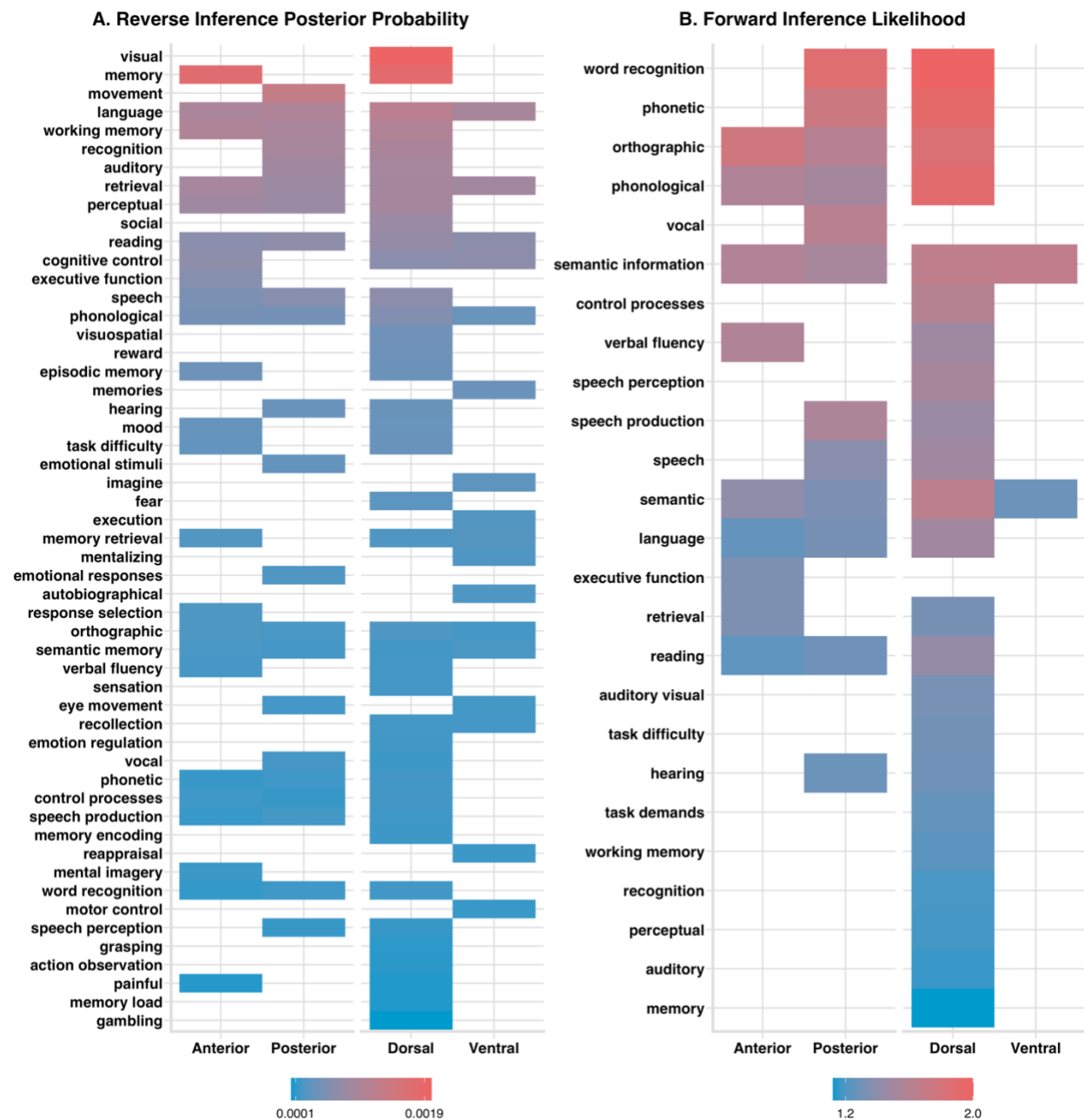
Figure S10. Functional terms associated with the IFG clusters derived based on the task-constrained gradients according to the A) specificity/reverse inference analyses and B) consistency/forward inference analyses. The colour indicates the effect sizes, with red colours suggesting greater association. Only statistically significant associations are highlighted. Synonymous terms with similar pattern of associations across the LIFG clusters were excluded.
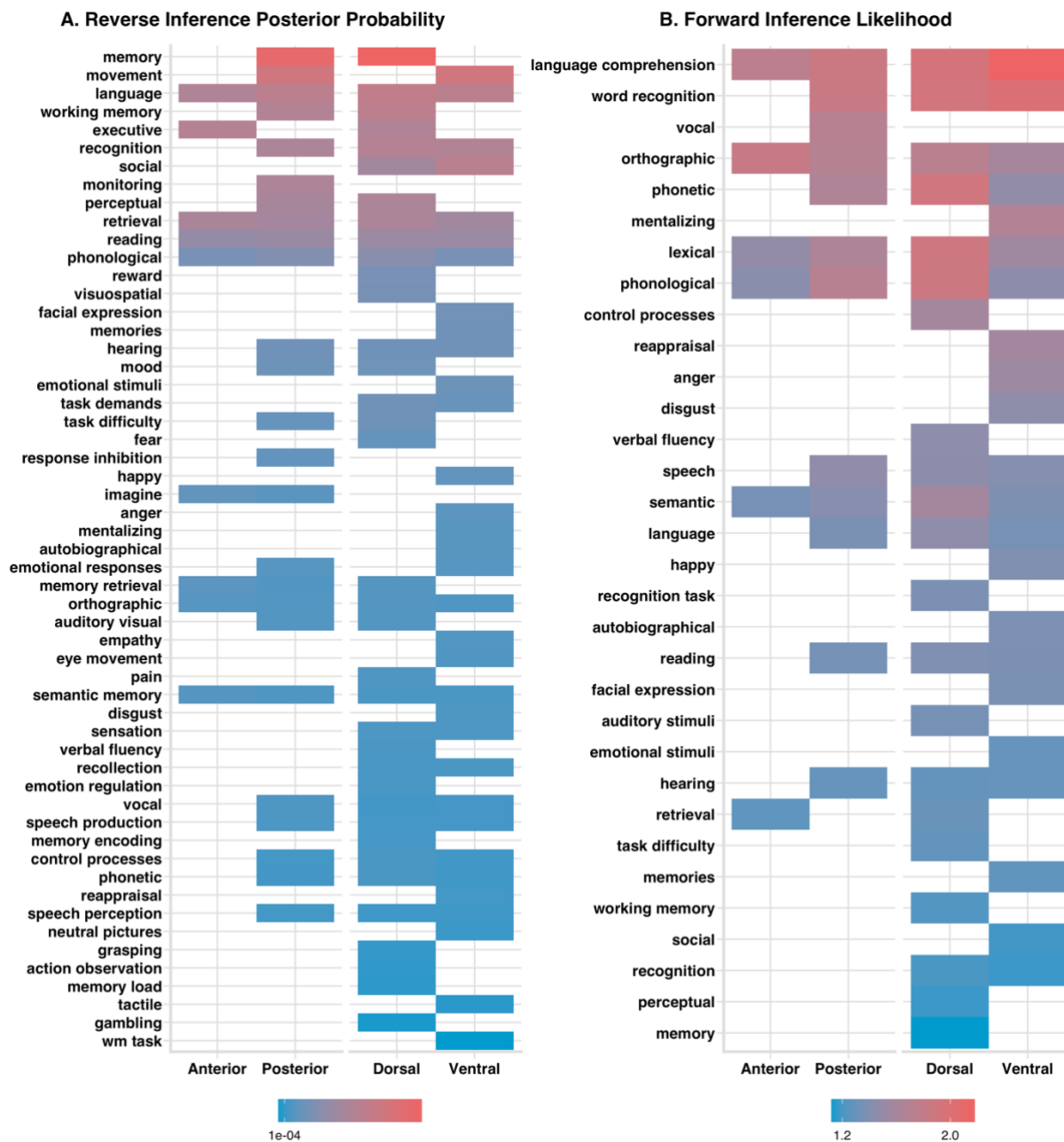
Figure S11. Functional terms associated with the IFG clusters derived based on the task-free gradients according to the A) specificity/reverse inference analyses and B) consistency/forward inference analyses. The colour indicates the effect sizes, with red colours suggesting greater association. Only statistically significant associations are highlighted. Synonymous terms with similar pattern of associations across the LIFG clusters were excluded.

Table S1. The MNI coordinates for the center of gravity of the hard clusters representing the edges of the task-free and task-based gradient maps.

Cluster	Task-free Gradients			Task-based Gradients		
	X	Y	Z	X	Y	Z
Anterior	-44	40	-6	-48	39	1
Posterior	-50	17	9	-44	21	2
Dorsal	-50	23	21	-51	21	22
Ventral	-44	27	-5	-44	38	-11

Table S2. The number of studies from the NeuroQuery database that reported at least one activation coordinate in each hard cluster. These studies were used as the input to MACM and functional decoding analyses.

Cluster	Task-free Gradients	Task-based Gradients
Anterior	664	851
Posterior	1064	1164
Dorsal	1332	1298
Ventral	1098	627

Table S3. Results of the seed-based resting-state functional connectivity analyses conducted on clusters extracted from the **anterior-posterior task-free gradient**.

Analysis	AAL Label	Cluster Size (mm^3)	Max Z Value	X	Y	Z
Anterior Cluster > Posterior Cluster	Frontal_Inf_Orb_L	51,200	21	-36	38	-12
	Angular_L	18,712	18	-44	-68	44
	Frontal_Inf_Orb_R	6,832	17	38	38	-12
	Cerebelum_Crus2_R	17,824	16	46	-72	-40
	Temporal_Mid_L	16,336	16	-64	-44	-10
	Angular_R	6,848	15	40	-72	46
	Cerebelum_Crus2_L	10,336	14	-42	-72	-40
	Temporal_Mid_R	6,568	14	62	-40	-8
	Rectus_R	4,216	14	4	44	-16
	Frontal_Inf_Tri_L	912	13	-52	24	28
	Frontal_Mid_R	2,952	11	32	20	50
	Cerebelum_9_R	520	11	2	-56	-50
	Precentral_L	672	11	-46	8	36
	Frontal_Inf_Tri_R	2,320	9	52	34	22
	Vermis_10	512	8	2	-48	-34
	Temporal_Pole_Mid_L	1,000	7	-32	8	-38
Posterior Cluster > Anterior Cluster	Frontal_Inf_Oper_L	17,720	20	-52	16	2
	Cingulum_Mid_L	35,744	19	-6	14	38
	Frontal_Inf_Orb_R	17,056	18	50	18	-4
	SupraMarginal_L	21,256	16	-56	-40	26
	SupraMarginal_R	14,000	16	58	-30	32
	Precentral_R	2,032	15	54	6	42
	Frontal_Mid_L	5,776	14	-30	50	24
	Frontal_Mid_R	2,760	14	34	46	30
Anterior Cluster ∩ Posterior Cluster	Frontal_Inf_Tri_L	26,192	22	-54	20	20
	Frontal_Inf_Tri_L	5,216	22	-54	22	4
	Supp_Motor_Area_L	23,232	19	-2	22	60
	Parietal_Inf_L	32,240	18	-54	-44	48
	Frontal_Inf_Tri_R	4,320	14	48	40	0
	Frontal_Inf_Orb_R	1,120	14	56	30	-2
	Frontal_Inf_Tri_R	1,704	13	58	26	18
	Cerebelum_Crus1_R	832	11	14	-76	-30

Temporal_Inf_L	2,424	9	-44	-2	-44
Temporal_Inf_L	856	8	-50	-6	-40
Fusiform_L	608	8	-44	-40	-18

Table S4. Results of the seed-based resting-state functional connectivity analyses conducted on clusters extracted from the **dorsal-ventral task-free gradient**.

Analysis	AAL Label	Cluster Size (mm^3)	Max Z Value	X	Y	Z
Dorsal Cluster > Ventral Cluster	Frontal_Inf_Tri_L	24,592	19	-44	32	18
	Parietal_Sup_L	26,328	18	-26	-70	46
	Precentral_L	968	16	-44	2	22
	Temporal_Inf_L	14,032	16	-54	-60	-14
	Frontal_Inf_Tri_R	10,872	15	50	38	18
	Frontal_Mid_L	5,280	13	-28	10	66
	Temporal_Inf_R	3,280	13	60	-50	-8
	Cerebelum_8_R	3,968	12	28	-70	-46
	Cerebelum_Crus1_R	832	10	6	-80	-24
	Insula_L	624	8	-42	-2	6
Ventral Cluster > Dorsal Cluster	Frontal_Inf_Tri_L	14,984	19	-42	26	0
	Frontal_Sup_Medial_L	58,464	16	-4	52	16
	Insula_R	12,960	14	32	20	-14
	Temporal_Inf_L	13,808	14	-48	2	-36
	Temporal_Mid_L	10,728	13	-58	-18	-10
	Temporal_Inf_R	9,960	13	48	2	-32
	Cerebelum_Crus1_R	2,232	12	28	-78	-32
	Temporal_Mid_R	4,344	12	54	-28	-8
	Angular_L	12,208	11	-54	-60	30
	Precuneus_L	1,576	11	-12	-52	32
	Temporal_Sup_R	712	10	58	-44	24
	Cerebelum_Crus2_R	648	9	24	-88	-38
Dorsal Cluster ∩ Ventral Cluster	Frontal_Inf_Tri_L	85,432	24	-54	22	18
	Frontal_Sup_Medial_L	21,624	17	-2	38	46
	Frontal_Inf_Tri_R	6,032	14	56	28	20
	Cerebelum_Crus1_R	7,016	14	14	-80	-30
	Temporal_Mid_R	5,424	8	54	-38	8
	Fusiform_L	5,272	8	-30	2	-44

Table S5. Results of the meta-analytic co-activation analyses conducted on clusters extracted from the **anterior-posterior task-based gradient**.

Analysis	AAL Label	Cluster Size (mm^3)	Max Z Value	X	Y	Z
Anterior Cluster > Posterior Cluster	Frontal_Inf_Tri_R	3,208	NA	48	36	18
	Occipital_Mid_R	5,432	4	34	-68	34
	Parietal_Inf_L	9,816	4	-38	-44	42
	Frontal_Inf_Tri_L	20,392	NA	-46	34	12
Posterior Cluster > Anterior Cluster	Insula_R	8,200	NA	44	18	-2
	Insula_L	22,616	NA	-42	18	0
Anterior Cluster ∩ Posterior Cluster	Frontal_Inf_Oper_R	18,200	4	44	20	10
	Supp_Motor_Area_L	15,224	NA	-2	18	46
	Pallidum_L	1,648	NA	-14	6	4
	Parietal_Inf_L	6,544	NA	-36	-54	44
	Frontal_Inf_Oper_L	41,856	NA	-42	18	14
	Occipital_Inf_L	2,312	NA	-44	-60	-12
	Temporal_Mid_L	1,208	NA	-56	-40	0

Table S6. Results of the meta-analytic co-activation analyses conducted on clusters extracted from the **dorsal-ventral task-based gradient**.

Analysis	AAL Label	Cluster Size (mm^3)	Max Z Value	X	Y	Z
Dorsal Cluster > Ventral Cluster	Rolandic_Oper_R	11,904	4	44	4	20
	Angular_R	2,408	4	32	-56	48
	Supp_Motor_Area_L	4,976	NA	-2	12	48
	Parietal_Sup_L	8,552	4	-24	-62	50
	Frontal_Inf_Oper_L	28,584	NA	-46	16	22
	Fusiform_L	520	3	-42	-58	-18
Ventral Cluster > Dorsal Cluster	Frontal_Inf_Orb_R	4,048	4	36	22	-18
	Frontal_Sup_Medial_L	1,080	4	-6	42	40
	Frontal_Inf_Orb_L	17,152	NA	-40	34	-10
	Angular_L	2,720	NA	-48	-62	30
Dorsal Cluster ∩ Ventral Cluster	Insula_R	9,136	NA	40	22	-4
	Supp_Motor_Area_L	9,800	NA	-2	20	46
	Frontal_Inf_Tri_L	30,888	NA	-44	20	12

Parietal_Inf_L	2,696	NA	-36	-56	46
Temporal_Mid_L	2,824	NA	-56	-38	0

References

Jackson, R. L., Hoffman, P., Pobric, G., & Lambon Ralph, M. A. (2016). The Semantic Network at Work and Rest: Differential Connectivity of Anterior Temporal Lobe Subregions. *The Journal of Neuroscience*, *36*(5), 1490–1501. https://doi.org/10.1523/JNEUROSCI.2999-15.2016

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fisch, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. https://doi.org/10.1152/jn.00338.2011