



Network Level Representation of Conceptual Content

CiMeC – Centre for Mind/Brain Sciences

Doctoral School in Cognitive and Brain Sciences (XXXI Cycle) Track Cognitive

Neuroscience

Candidate: Aidas Aglinskas

Advisor: Prof. Scott Fairhall

TABLE OF CONTENTS

ABSTRACT	4
CHAPTER 1: GENERAL INTRODUCTION	7
KNOWLEDGE IN THE BRAIN & PREVIOUS WORK.....	7
THE ORGANISATION OF CONCEPTUAL KNOWLEDGE IN THE BRAIN – THEORETICAL PERSPECTIVES	8
EMBODIED THEORIES	8
DISEMBODIED COGNITION	9
THESIS STRUCTURE	14
CHAPTER 2: BEHAVIORAL MEASURES OF PERSON KNOWLEDGE.....	17
INTRODUCTION	17
METHODS	20
RESULTS	25
DISCUSSION.....	30
CHAPTER 3. REGIONAL SPECIALISATION AND COORDINATION WITHIN THE NETWORK FOR PERCEIVING AND KNOWING ABOUT OTHERS	34
INTRODUCTION	34
DATA ACQUISITION	44
DATA ANALYSIS.....	46
REGIONAL TUNING OF ATL	64
COGNITIVE TAXONOMY IN PERSON KNOWLEDGE.....	66
CHAPTER 4: ORGANISATION OF PERSON KNOWLEDGE ACROSS STIMULUS MODALITIES	70
INTRODUCTION	70
RESULTS	74
DISCUSSION.....	80
CHAPTER 5: CONCEPTUAL KNOWLEDGE OF OBJECT CATEGORIES	84
INTRODUCTION	85
METHOD	88
RESULTS	96
CHAPTER 6: GENERAL DISCUSSION.....	110

Abstract

Our ability to store knowledge and represent the world within our minds has spanned multiple disciplines (philosophy, psychology, neuroscience). Currently, theories of human conceptual knowledge suggest that human representation of the world is widely distributed across the brain. Regions involved in sensory/motor simulation as well as amodal systems contribute to our flexible ability to manipulate knowledge. A detailed account of how the overall human semantic system works at a network level is still lacking.

To begin our investigation into how knowledge is distributed across brain networks, we will first consider a specific kind of knowledge - person related knowledge. Chapter 2 will look at the behavioural indicators of person-knowledge organisation. We will ask participants to judge explicit/subjective similarity between different person-knowledge domains: social, physical, biographical, episodic and nominal knowledge. This will allow us to investigate whether these processes are independent or related to each other. We will then compare these judgements to implicit similarity measures to see whether correlated patterns of responses or reaction are informative about cognitive similarity. Chapter 3 will look at how the brains core/extended system for face perception coordinates across the aforementioned person-knowledge domains. We will investigate the representational similarity of different person-knowledge domains in individual regions, and crucially - across the network as a whole. This will allow us to address whether cognitions are localised in individual regions or distributed across the whole network. Chapter 4 will investigate the stability of network organisation when going across modalities. Extended system for face perception has been shown to be recruited during

familiar name reading. We will ask whether network-level patterns of activation during person-knowledge remain stable across input modalities. Chapter 5 will generalize the network-level approach to investigate broader semantic categories. We will interrogate how diverse regions activated during semantic processing, interact during processing of naturally occurring conceptual categories. We will use a corpus derived semantic distance model and compare it to individual region activity to that of the network overall. We will ask whether information about conceptual distance between categories is contained within individual regions or arises as a product of coordinated effort across the network.

Combined, evidence presented in this thesis speak to the distributed nature of cognitive representation. Different kinds of person-knowledge and object categories are highly linked and rely on overlapping neural substrates. We demonstrate that instead of being specialised for particular tasks, brain areas involved in meaning extraction tend to be involved in most kinds of conceptual processing. Individually regions have slight cognitive tunings and can be geared towards specific cognitions. Differences in person-knowledge and object categories emerge as a product of the coordinated interplay between multiple brain regions.

Acknowledgements

Many thanks to Scott Fairhall for the incredible mentorship and patience.

Additional thanks to my colleagues within FAIRLAB and CIMeC.

Chapter 1: General Introduction

Knowledge in the brain & Previous work

Our ability to store knowledge and represent the world within our minds has spanned multiple disciplines: philosophy, psychology & neuroscience. Conceptual knowledge refers to our ability to represent objects and ideas that are abstracted away from individual exemplars or experiences. Our ability to store, manipulate and combine complex concepts differentiates us from other animals and has been the basis for incredible human achievements of technology, culture and arts. How the brain represents knowledge is currently unclear. Theories investigating the mechanisms enabling our capacity for conceptual knowledge suggest that human representation of the world is widely distributed across the brain. It also is, at least in part, shaped by and depends on the senses through which humans acquired that knowledge (Barsalou, Kyle Simmons, Barbey, & Wilson, 2003). Further studies support a view of a “lumpy” semantic space, - knowledge of the world is represented based on attributes such as manipulability size & color (Thompson-Schill, 2003). Recent investigations have shown that tools - a recent invention in evolutionary terms, activate a common brain area (Bracci, Cavina-Pratesi, Ietswaart, Caramazza, & Peelen, 2012) – suggesting that concepts belonging to the same conceptual categories might be stored in overlapping brain regions. However, a detailed account of how the overall semantic system works is still lacking. Several prominent theories have been proposed to describe how knowledge might be organised in the brain.

The organisation of conceptual knowledge in the brain – theoretical perspectives

Questions of how the brain represents various concepts has been sparked by early work with brain damaged patients, when researchers observed that damage to specific brain areas produced impairments in recognising very specific object categories. Selective impairments of categories such as people, tools, animals, plants, musical instruments & buildings have been observed (Caramazza & Shelton, 1998; Warrington & Shallice, 1984), - suggesting that semantic knowledge may be organised in lumps - knowledge about exemplars of a particular category are stored together in the brain. While being influential, this account does not account for concepts that cross category boundaries. Whether our knowledge of the meaning of the word “handyman” is located in a single region or depends on our knowledge of both people and tools - is unclear. Several other theories have been proposed to elucidate how different sources of knowledge get integrated to form conceptual knowledge.

Embodied theories

Embodied theories of cognition suggest that modalities through which we acquire a concept, play a major role in understanding that concept. Under this view – retrieving a concept involves active simulation of our motor and sensory experiences and the interactions between modality specific neural populations (Barsalou, Simmons, Barbey, Wilson, 2003; Tyler, Moss, Durrant-Peatfield & Levy, 2000). Evidence from both neuropsychology and neuroimaging provide some evidence for this theory. Perceiving and knowing about colour relies on shared neural substrates (Simmons et al., 2007). Thinking of motor related action concepts, activates primary motor cortices (Hauk,

Johnsrude & Pulvermüller, 2004; Papeo et al., 2015). Recalling perceptual knowledge activates visual cortices (Goldberg, Perfetti & Schneider, 2006). Additional evidence suggests that patients with motor disorders, have trouble with knowledge related to motor concepts, reinforcing the link between sensory-motor cortices and concept knowledge. However, embodied frameworks do not explain the organisation of concepts that rely little on sensory-motor features, such as abstract concepts. Some accounts therefore stress the importance of high-level, abstract representations of conceptual knowledge to account for the versatility of the conceptual system (Caramazza, Anzellotti, Strnad & Lingnau, 2014).

Disembodied cognition

In contrast, disembodied theories suggest that conceptual knowledge representation resides outside the sensory and motor cortices, and any activation in sensory-motor systems is epiphenomenal. Disembodied cognition accounts stress that conceptual knowledge is located within amodal regions (Mahon & Caramazza, 2008). Evidence showing dissociation between concept and feature knowledge support this theory. Mahon & Caramazza (2005) demonstrated that patients can have impairments in using a tool, but can still name it, inconsistent with embodied theories. These theories suggest an amodal, symbolic representation of conceptual knowledge. However, selective impairment in knowledge about concepts related to motion in patients with Parkinson's disease (Boulenger, et al., 2008), or when motor cortex is stimulated with TMS (Friedemann Pulvermüller, Hauk, Nikulin & Ilmoniemi, 2005) suggest an active, rather than an epiphenomenal role for motor areas in semantic knowledge. Most current theories

acknowledge that both sensory-motor cortices and amodal regions (or systems) participate in semantic knowledge representation (Binder & Desai, 2011; Patterson & Lambon-Ralph, 2016; Mahon & Caramazza, 2008).

Distributed plus hub

Aligning embodied and disembodied theories, Patterson, Nestor & Rogers (2007) proposed a distributed, hub-and-spoke model. At its core, the model suggests an interaction of both embodied and amodal (disembodied) systems. Modality specific stores ('spokes') contribute information to a single 'hub' where generalised, amodal representations emerge. This model provides more information in explaining conceptual processing and its disorders, such as the graded deterioration of knowledge in semantic dementia than embodied or disembodied theories in isolation. Originally proposed with a single hub, located in the bilateral ATL, this theory was later expanded to allow for two (ATL and AG) hubs (Lambon-Ralph, Jefferies, Patterson & Rogers, 2016). Accumulating evidence, however suggest the possible existence of more areas that receive abstracted, multimodal input (Binder et al., 2009). These results raise the possibility of a network of hub-like regions (including but not limited to ATL) integrating broad kinds of semantic information.

The Internalised semantic system

Studies contrasting semantic versus non-semantic content (meaningful v. meaningless words, familiar vs. unfamiliar proper nouns, famous vs. unknown people) frequently

observe activations in a stable collection of regions spanning parietal, temporal & frontal lobes. In meta-analysis, Binder et al (2009) described a set of regions, active for a variety of meaningful stimuli. These regions have been termed *internalised cognition regions* and are involved in meaning extraction regardless of stimulus category (Binder et al, 2009) Anatomically, these regions reside outside of sensory-motor cortices, suggesting an integrative or representational, rather than perceptual function. Moreover, Fairhall and Caramazza, (2013) observed cross-modal representations were limited to these internalised regions, further emphasizing their importance in non-perceptual, conceptual representation.

Multiple ‘hubs’

Regions associated with the internalised semantic processes, described by Binder et al., (2009) raise the possibility of multiple semantic, disembodied ‘hubs’. A set of regions located outside of sensory cortices, and active during a variety of semantic tasks suggest their role general meaning processing. Evidence reviewed by Pulvermüller (2013) suggest that multiple brain regions receive multi-modal input from sensory-motor cortices suggesting a distributed set of binding sites, instead of the single one proposed by Patterson et al. (2006). Binder and Desai (2011) suggest that multiple regions across the brain bind information from different sensory-motor cortices and that this network of high-level semantic areas jointly participates in complex semantic tasks. Functional roles of the regions comprising the internalised semantic system - whether they perform

complementary or redundant functions is unclear. This is due, at least in part – because these regions are co-activated across many different cognitive efforts (see next section).

Diverse functions

Striking feature of the brains internalised cognition system is that it can be observed activated not only across different semantic categories and modalities - from verbs, nouns, objects and faces, - but also across vastly different cognitive efforts. Parts of the internalised cognition network participate in autobiographical memory (thinking about the self) (Hassabis & Maguire, 2007, Schacter & Addis, 2007), theory of mind (thinking about others) (Schilbach et al, 2008), trait attribution (social cognition) (Frith, 2007), episodic memory recall & prospection. The common feature of these vastly different cognitive efforts is the need to retrieve and manipulate conceptual content. This suggests that the brains semantic system is involved in diverse kinds of conceptual processes. The semantic network is largely identical to the default-mode-network (DMN), originally characterised by task induced de-activation (Raichle et al., 2000). Activation of DMN during rest has been linked to mind-wandering - incidental retrieval of semantics and episodic memories when the mind is at rest. Components of this networks remain active even during deep sleep (Horowitz et al., 2009). The extreme versatility of the brain internalised cognition network makes it difficult to study - if a brain region (or a collection of them) is activated for everything - how do we know what it does? One feasible starting point could be to reduce the complexity by considering a single stimulus category.

Single Category: People

Conspecifics represent a rich source of diverse information. Each individual person has with them associated diverse kinds of information. Unique episodic memories, biographical facts, names, and social traits can be accessed by looking at someone we know. By considering a single stimulus class, it's possible to gain a deeper, more precise understanding of knowledge organisation. People are a class of unique stimuli, and person's identity can be accessed by different modalities, such as by looking at their face, reading their name or listening to their voice. At a neural level, human faces elicit activity in both perceptual and semantic areas (Haxby, Hoffman, Gobbini, 2000) - making them a suitable stimulus to investigate the interaction between the two systems. Two important aspects that might help shed light on it are investigation of representational content within regions and functional coordination between network elements.

Co-ordination vs co-activation

Do functionally co-activated regions perform similar or different functions? To answer this, we need to compare representational spaces of regions involved. Approaches investigating representational spaces have been informative in untangling the processing differences in co-activated regions. Fairhall & Caramazza (2013) showed that within a network of regions active during semantic processing, only a handful were sensitive to semantic distance, as indicated by comparing representational spaces to a semantic model. Representational similarity analysis (RSA) is a pivotal tool in elucidating regional function. RSA abstracts from activation magnitudes, instead focusing on similarity in

patterns of activity. Using representational similarity matrices (RDMs) as a multivariate measurement unit, RSA enables us to computationally test what stimulus dimensions/features a particular brain region represents. For example, to investigate whether a particular brain region is sensitive to visual similarity or semantic category, one can fit competing models of expected similarity and then test them against the fMRI data. By comparing representational spaces between two or more brain regions, we can infer regional co-ordination. Among many regions co-activated to perform a function (such as face perception), regions co-ordinating to perform a shared or similar function would be expected to share response pattern similarity. Castello, Halchenko, Guntupalli, Gors & Gobbini (2017) demonstrated the efficacy of this procedure (second order RSA – similarity of similarity spaces). By comparing multivariate response patterns of regions involved in face perception authors found that groups of regions traditionally associated with perceptual functions - early visual cortex, occipital and fusiform face areas, grouped largely separately from regions associated with cognition – Precuneus, TPJ and mPFC.

Thesis structure

I will begin this thesis by investigating what information about the relationships between different kinds of person-knowledge can be learned from behavioural responses (Chapter 2). By utilising multivariate analysis techniques (clustering, RSA) we investigate whether patterns of behavioral responses contain information about underlying structure of person-knowledge. Chapter 3 will extend the approach used in Chapter 2 to investigate the neural correlates of person-knowledge. We will interrogate the coordinated activity of

core/extended person-knowledge systems. By looking at network level activity we will look at person-knowledge domain similarity and regional coordination. This will allow us to answer whether different kinds of person-knowledge that are similar behavioral, are similarly represented in the brain. We look for groups of regions, sharing multivariate response patterns to infer regional co-ordination in the large coactivated network for person knowledge. Chapter 4 will investigate how the extended component of the person-perception system behaves when person-knowledge is accessed via written names. This will establish whether relationships between different person-knowledge domains are perceptually driven or represent a stable feature knowledge encoding. Chapter 5 will generalize the approach to broader object categories. We will investigate whether knowledge about conceptual distance is localized in individual regions or distributed across a broader network. By utilizing corpus derived conceptual distance model we will contrast whether there is more information about conceptual distance located in individual regions, or whether this information arises at the network level. This direct comparison of local versus shared information will allow us to quantify the degree to which conceptual knowledge is distributed across brain networks, as opposed to localized in individual regions.

Concisely, this thesis will present evidence for distributed knowledge representation. In a series of experiments utilizing multivariate measures, we will look at what information structure is represented across brain networks. We will show that conceptual relationships between different person-knowledge domains, as well as object categories are visible in the network-level patterns of regional interactions.

Chapter 2: Behavioral measures of person knowledge

Introduction

Knowledge is broad and wide-ranging. The richness of this repertoire is particularly evident in our knowledge about other people. Person knowledge includes classical semantic knowledge such as a person's occupation. It also includes knowledge about their appearance, episodic memories, nominal knowledge as well as knowledge about personality traits such as friendliness and trustworthiness (social knowledge). Transient impairments evident in situations where we forget someone's name or feel like we've met someone but fail to recognize them suggest different types of person knowledge is at least somewhat independent. Bruce and Young's (1986) model includes different nodes for naming and person recognition to explain these phenomena. Other accounts argue for a correlated structure of person knowledge (Wiggins, 1979). Frequently observed, impairments in recalling episodic memories associated with a person are correlated with impairment to factual knowledge retrieval, suggesting at least some degree of correlation (Seidenberg et al., 2009). Todorov, Said, Engell & Oosterhof (2008) demonstrated that people differ along multiple, highly related social dimensions, such as attractiveness, trustworthiness, dominance & threat. The subjective difference between cognitions, and no associated measurement units makes building a cognitive ontology difficult - how can we compare attractiveness of a person to the memories associated with them?

Measuring things that have no associated units or measures is challenging. A common measure used to combat this is multivariate similarity. Similarity and its inverse - discriminability are the bases for many modern multivariate tools, such as hierarchical clustering analysis and representational similarity analysis (RSA).

Hierarchical Clustering

Clustering analyses reveal hierarchical patterns in similarity structures, visualized by dendrograms. Frequently used in taxonomic research, clustering arranges multivariate features, such as phenotypic trait descriptions of different animals (e.g. has claws, is small) to form cohesive groups. In this instance, hierarchical structure of the dendrogram reflects increasingly superordinate categories, from felines to mammals to vertebrates.

RSA

Representational similarity analysis (RSA) is a framework for testing similarity-based hypotheses. By constructing a model of expected pattern of similarities, it is possible to evaluate information content derived from heterogeneous sources. RSA facilitates comparing different measures (Cichy, Pantazis & Oliva, 2014), neuroimaging data across species (Kriegeskorte, 2009) and comparing representational spaces between humans and artificial neural networks (Cichy, Khosla, Pantazis & Oliva, 2017).

Task selection

We have chosen to inspect broad ranging kinds of person-knowledge. Haxby et al (2001) and Gobbini & Haxby (2007) models of person knowledge suggested that Social, Episodic and Biographical knowledge are represented (at least partly) in different brain regions. Additionally, we decided to investigate two further elements of knowledge about

other people. We investigate how accessing predominantly physical knowledge, such as attractiveness and distinctiveness of a person's face relates to other kinds of person-knowledge. Lastly - retrieval of proper names (nominal knowledge) has been suggested by Bruce and Young (1986) to be a distinct process than recalling other kinds of facts and was included as a separate category.

Divisions between domains are sometimes soft boundaries – some person knowledge domains are similar and interlinked. Responses of traditionally perceptual judgements such as attractiveness and social judgements, like trustworthiness are correlated across individuals (García, Beltrán & Calvo, 2018). This raises a question whether these two tasks rely on the same cognitive mechanism. Neuroimaging evidence has shown that Attractiveness and Trustworthiness activate different brain areas. In a meta-analysis, Bzdok, Langner & Caspers (2011) demonstrated that judging facial attractiveness engaged brain areas associated with reward while trustworthiness preferentially engaged mentalizing regions of the brain.

Organisation of person knowledge

Harnessing recent advances in computational tools (clustering, RSA) opens up new avenues for cognitive science. By investigating how people differ across multiple person knowledge dimensions, we can start to build a data driven cognitive ontology.

In this first experimental chapter, we report an investigation of the behavioral indicators of the relationship between different types of person-knowledge domains. We start by asking participants to provide person-knowledge ratings for 40 famous people. We first

model subjective similarity by explicitly asking participants which tasks they thought were similar. We then compare it to different, implicit behavioural indicators of cognitive processing. We do this by separately clustering reaction times, as well as person knowledge ratings. We then employ RSA to elucidate which measures provide more information about cognitive structure.

Methods

Subjects

25 subjects underwent fMRI scanning (see chapter 3), during which they looked at pictures of famous people. Participants answered 10 separate person specific questions (see *experimental conditions* below). Subjects indicated their responses via button boxes provided (4 buttons). fMRI data will be analysed separately in next chapter, current chapter will focus on concurrently collected behavioural measures (responses, reaction times and subjective task similarity ordering). 19 subjects also took part in the post-scanning, multiple-item-arrangement (MIA) task (see below: Explicit/Declarative Cognitive Similarity). In the current analysis we analyse the data from 19 subjects that participated in both person-rating (RT, response data) and MIA tasks (subjective similarity ratings).

Experimental conditions

During fMRI acquisition, participants looked at faces of 40 celebrities and answered ten different questions about each person. The questions were chosen to span five person-

knowledge domains: nominal, physical, social, episodic and biographical knowledge. For each domain, we selected two questions that require access to that domain (figure 1.1). Responses to nine questions ranged from 1 to 4 (ordinal scale), one question ('Occupation') had predefined answer categories. For more detailed description of the task, see chapter 3.

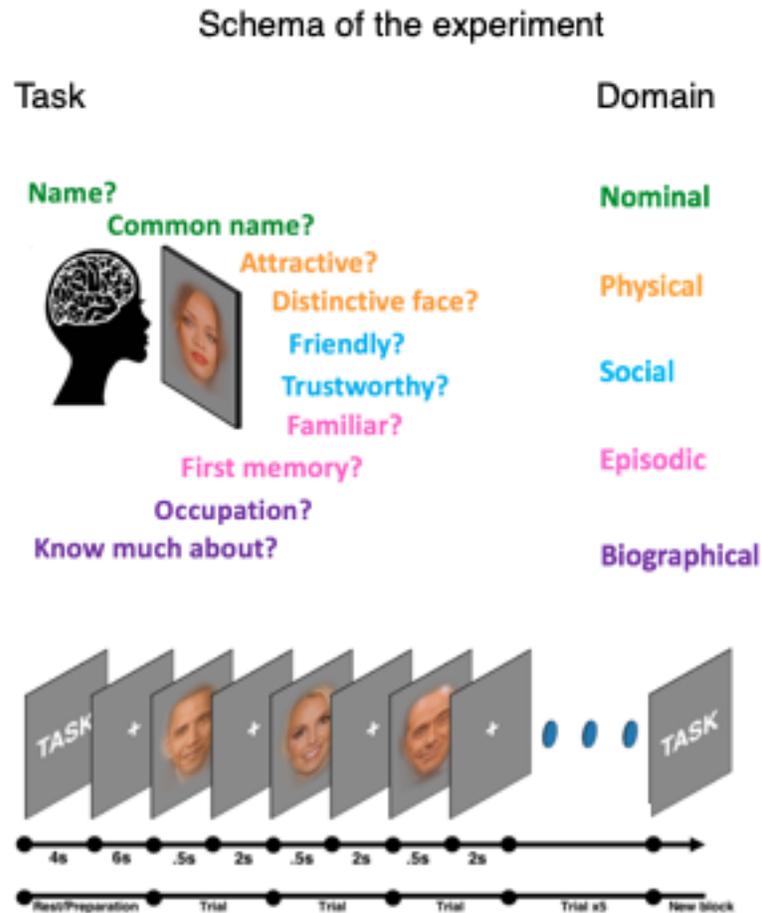


Figure 1.1. Experimental Conditions and procedure. Top) For each of the five cognitive domains (e.g. nominal knowledge) we included two related questions (“How well do you know their name”, “how common is their first name”). Tasks are colour coded according to the domain they were sampled from. Bottom) Each block was preceded by the task

name, faces were presented for .5s, followed by 1.5s of fixation cross, during which participants provided their response.

Explicit/Declarative Cognitive Similarity (Multiple Item Arrangement, MIA)

Immediately after the fMRI experiment, participants took part in post-experiment task, in which they were asked to arrange tasks they just performed according to their similarity. They were presented with a circular arena and asked to drag similar questions closer together, procedure known as multi-item-arrangement (MIA). Throughout the procedure subjects iteratively arrange items in an arena presented on the screen according to their similarity (Kriekeskorte & Mur, 2012). After items are arranged, pairwise dissimilarity structure can be inferred from the two-dimensional distance information, inverse of multidimensional scaling (MDS) procedure. By utilizing this approach, we can estimate the subjective pairwise similarity of different cognitive tasks (figure 1.2).

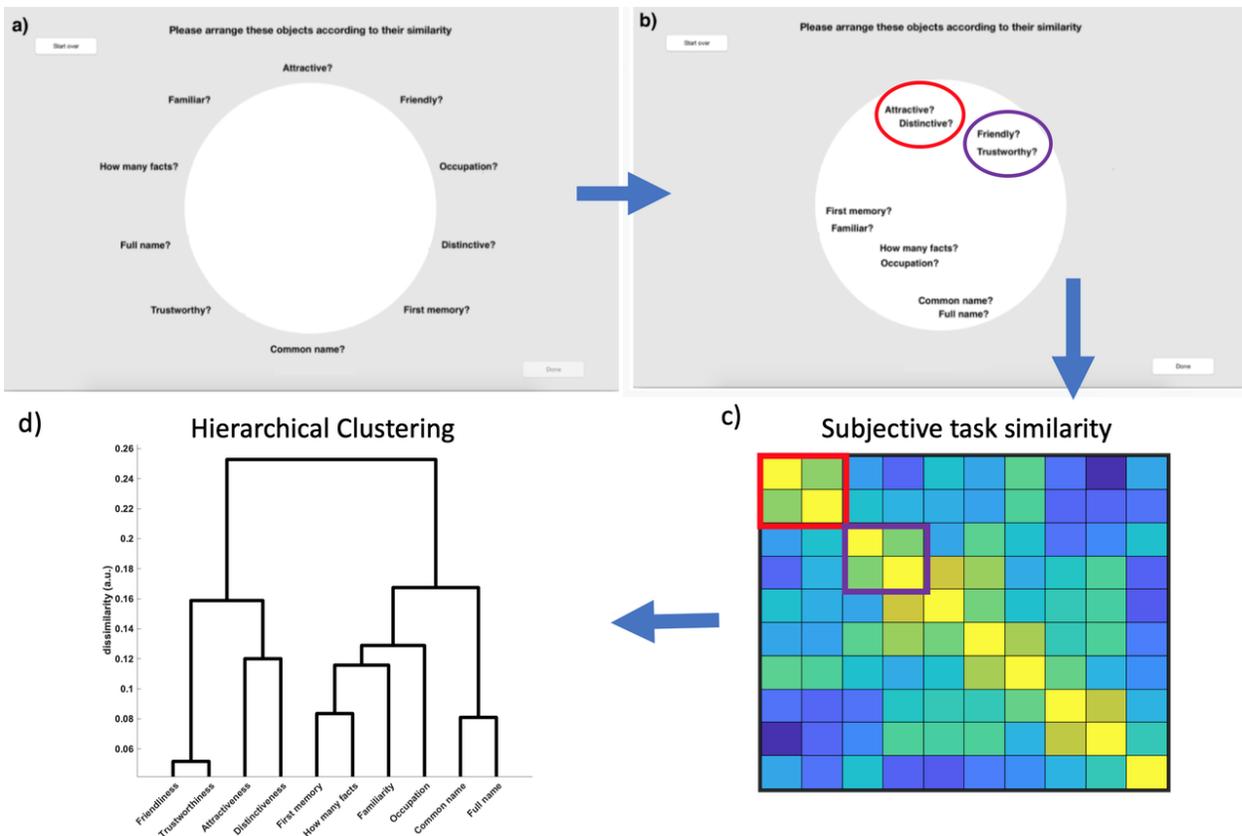


Figure 1.2. MIA procedure. A) Starting position of MIA, participants are asked to arrange tasks according to subjective similarity. B) Participants arrange tasks, placing similar tasks close together and dissimilar tasks further apart. C) Once done, pairwise similarity structure for ten tasks can be inferred. D) Correlation matrix is then clustered to reveal hierarchical relationships.

Response Similarity

Another way of calculating task similarity is to look for similar patterns of responses. For example, if people who are rated as more attractive are also rated as more trustworthy, - this correlated pattern of responses would suggest a commonality between the two tasks. Similarly, if friendliness and trustworthiness judgements rely on similar cognitive

mechanisms, we can expect reaction times when performing these two tasks to be more correlated to each other than to other tasks. We calculated pairwise similarities for both patterns of responses and patterns of reaction times, figure (1.3).

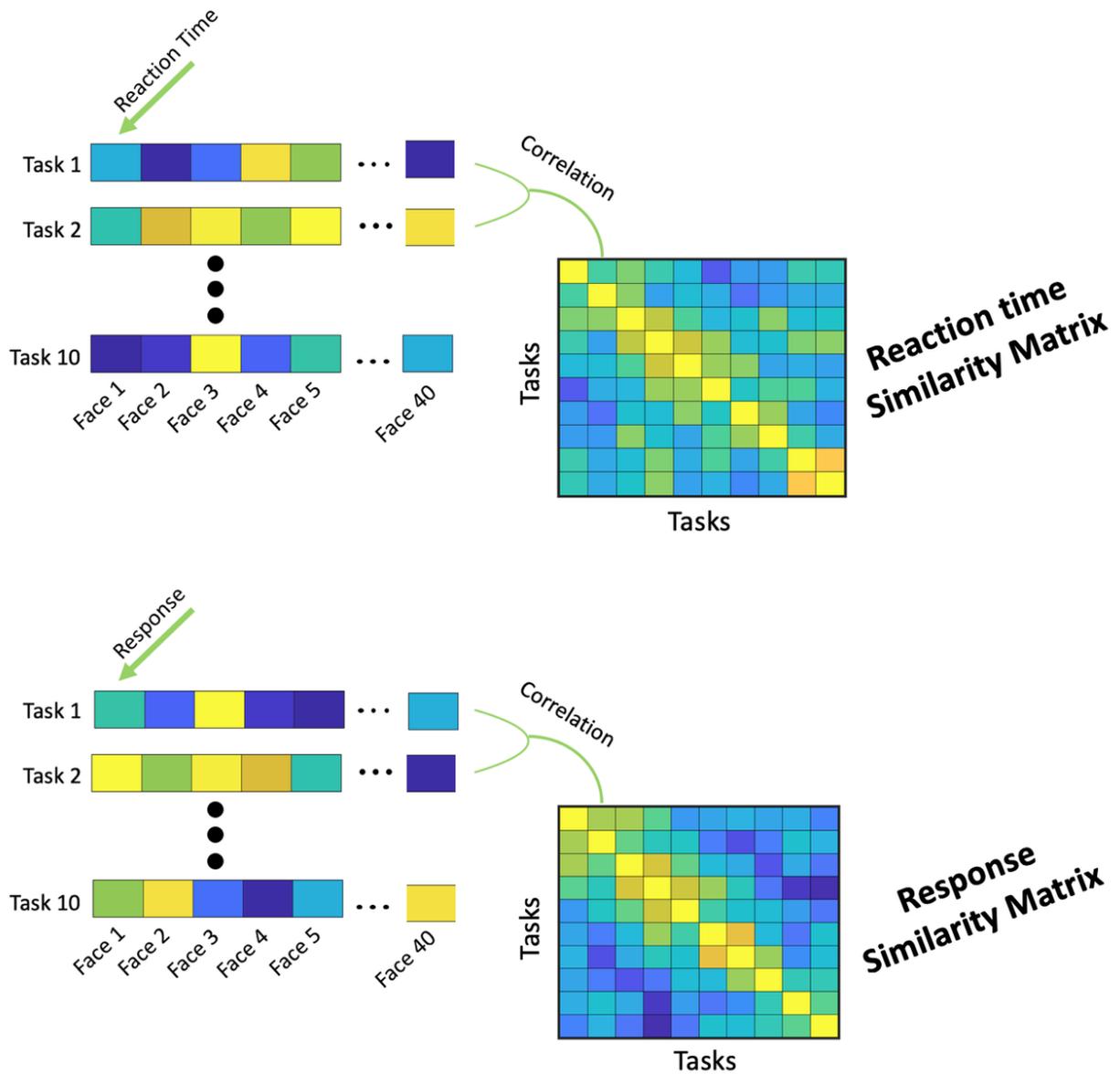


Figure 1.3. Implicit similarity measures. Patterns of responses across 40 faces were correlated for each pair of tasks. 10-by-10 correlation matrices representing response similarity and reaction time similarity were then subjected to further RSA analyses.

Results

Here we evaluated how well different behavioral measures (MIA, responses, RTs) capture task similarity relationships. Broadly, we investigate whether different person-knowledge domains are independent or related to each other. Similarity/independence of different tasks could imply shared/distinct cognitive mechanisms. Here we investigate both subjective and implicit behavioral measures related to person-knowledge similarity.

Visualizing MIA results in a group-average dendrogram demonstrates the correlated nature of person knowledge (Figure 1.4). Tasks requiring social and physical judgements were rated as being similar to each other, forming a cluster. They also appear distinct from biographical, episodic and nominal knowledge tasks, which form a separate cluster. Tasks selected from the same domain ('friendliness' and 'trustworthiness' tasks both belong to social domain of person knowledge) are often grouped close together. This hold true for social, physical & nominal knowledge, but not for biographical and episodic knowledge. Episodic and biographical judgements cluster together, distinct from nominal knowledge tasks. To see whether these results are common to all subjects, we next apply RSA measures.

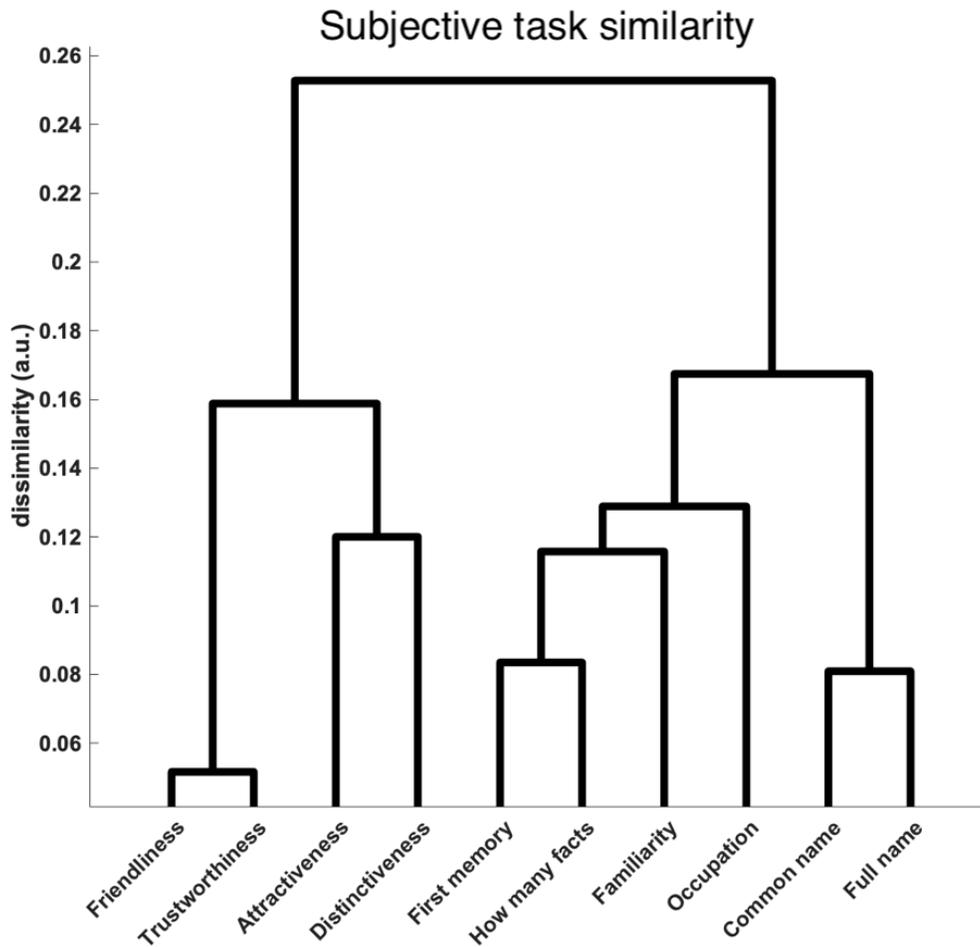


Figure 1.4. Subjective similarity ratings. Average MIA clustering of experimental tasks. Social and perceptual judgements cluster apart from biographical, episodic & nominal knowledge.

Next, we constructed an RSA task couplet model to test whether tasks sampled from the same domain were rated being more similar to each other than to other tasks. For example, we expect that rating how friendly someone is will be more similar to ratings of trustworthiness than to any other task in all measures (MIA, response and RT patterns), figure 1.5. Response scheme for occupation task did not follow ordinal scale, so we removed it and its biographical knowledge couplet ('how many facts do you know

about this person’) from RSA analyses. We will relate these biographical domain tasks to other person-knowledge exemplars on the bases of their neural signatures in chapter 3.

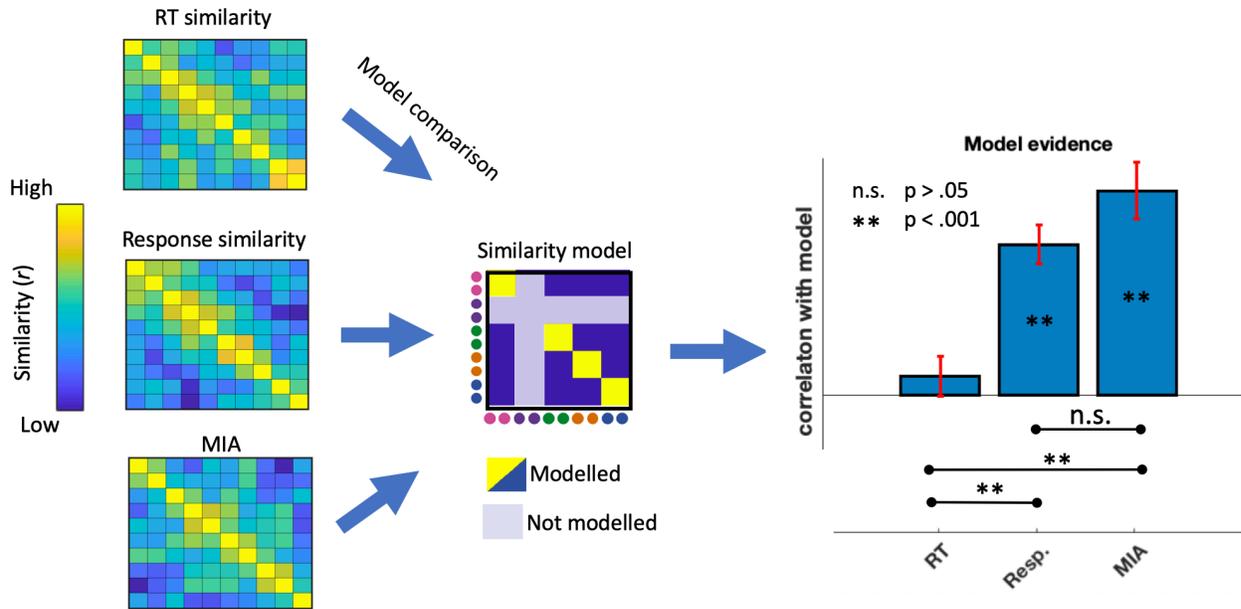


Figure 1.5. Data analysis pipeline and results. Similarity matrices derived from three different measurements - Reaction Times (RT), Response Ratings (resp), and subjective multiple-item arrangement task (MIA) were compared to an a priori expectation model (similarity model). Correlations between individual subjects’ similarity matrices and the model were then compared in a paired samples t-test (model evidence).

MIA similarity significantly correlated with the expected (task couplet) model $t(18) = 7.23, p < 0.001$; task sampled from the same domain (e.g. ‘friendliness’ and ‘trustworthiness’) were rated as being most similar. Within domain similarity is also evident from response pattern similarity. Significant similarity model fit revealed that

tasks from the same domain elicited strongly correlated patterns of responses $t(18) = 7.88, p < 0.001$. Reaction times did not correlate with the task couplet model $t(18) = 0.96, p = 0.352$, suggesting that RT similarity provides no information about task relatedness.

Next, we directly compare which measures provide more information about task relatedness. To do this we contrast the model fit of different measures in a paired samples t-test. This analysis revealed no difference between model fit of response patterns and MIA, $t(18) = 1.67, p = 0.111$; Indicating similar efficacy in explaining task similarity. When we correlate the two measures we observed that for each subject, their response pattern similarity matrix was correlated to their later MIA $t(18) = 4.25, p < 0.001$.

Both response similarity and MIA provided a better model fit than reaction times; $t(18) = 4.71, p < 0.001$ and $t(18) = 4.89, p < 0.001$ respectively. Individual subject's RT matrices did not correlate to their response pattern similarity matrix $t(18) = 1.57, p = 0.134$ or their MIA $t(18) = 0.68, p = 0.506$.

People clustering

Here we present a proof-of-concept grouping of people rated similarly across nine person-knowledge dimensions. Pairwise person similarity was calculated as their rating correlation across 9 cognitive tasks, figure (1.6).

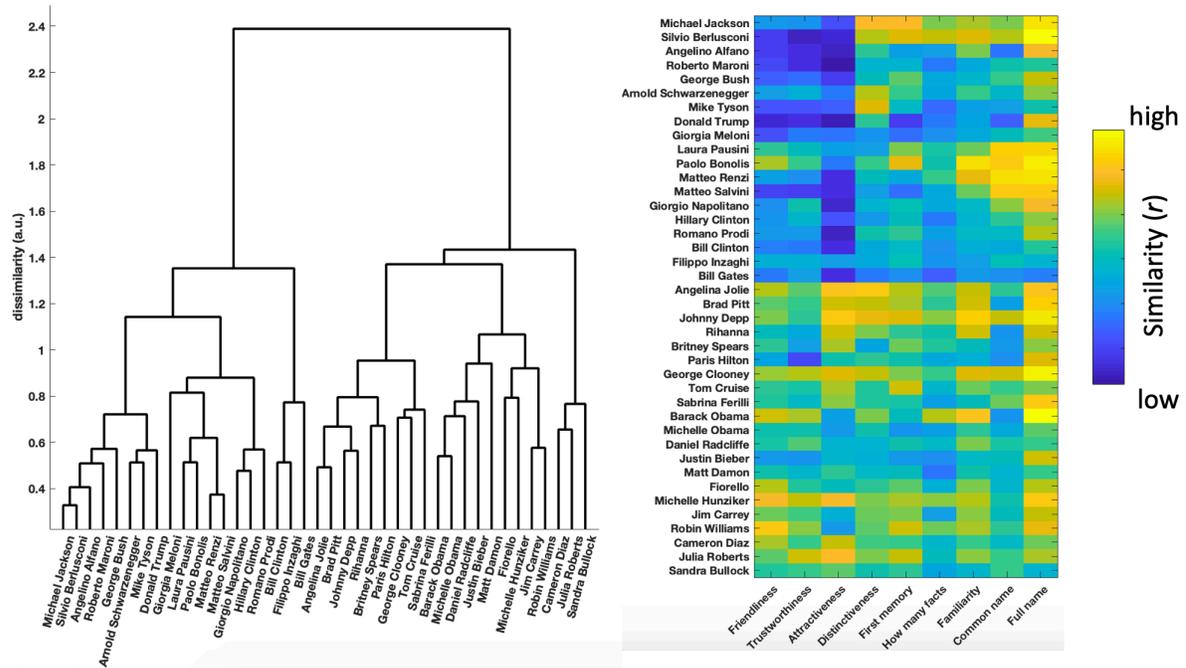


Figure 1.6. Famous person similarity across nine person-knowledge domains. Left) Clustering of different famous people reveals two prominent clusters. Right) Matrix of person-knowledge ratings, sorted by people and task similarity.

Visualizing the results, we see two prominent clusters. While there is no single feature (such as occupation or gender) describing the split, famous people grouped by the clustering procedure, tend to be rated similarly on multiple person-knowledge tasks. Looking into the responses arranged by task similarity, we observe that people rated low on friendliness, tend to be ascribed low scores in trustworthiness and attractiveness as well. This demonstrates the correlated nature of these cognitive efforts and offers a validation of the cognitive cluster describe above.

Discussion

In this chapter we investigated the behavioral indicators of the linked nature of person-knowledge. Establishing whether different knowledge domains are independent or related to each other is an important aspect of knowledge organisation. We observed that some forms of person-knowledge are more related than others. Converging evidence from both subjective task arrangement, and correlated patterns of behavioral responses suggest that social and physical knowledge domains might rely on related cognitive mechanisms. In the broad context of person knowledge spanning episodic and biographical domains of knowledge are rated as being more similar to each other, than they are to social, physical and nominal knowledge.

We used a similarity analysis approach to investigate knowledge of other people. By applying similarity measures to broad domains of person-knowledge we investigate efficacy of behavioral measures in uncovering cognitive structures. Using different behavioral metrics: reaction times, person knowledge ratings and subjective arrangement of task similarity we investigate the relatedness of cognitive processes. We sampled five domains of person knowledge - physical, social, biographical, episodic, and nominal. For each domain we selected two probe questions that tap into that domain. We used two different approaches, explicit ordering of tasks based on subjective similarity (MIA), and implicit patterns of response similarity (response similarity, RT similarity). These different measures can be compared via representation similarity analysis (RSA), we fit an a priori constructed model to see which measures capture expected task relationships.

MIA

By asking participants to arrange similar tasks, we investigated subjective task relationships. By inferring the pair-wise similarity space from a two-dimensional arrangement we obtain subjective hierarchy of task ordering. We first observed that task couplets tend to go together, validating our task selection. For social, physical and nominal knowledge domains, the task placed closest, was a task from the same domain. This suggests that we reliably sampled these domains of knowledge. At the next grouping level, we see relationship emerge not only between tasks, but also task domains. Social and perceptual judgements are rated as being distinct from other kinds of knowledge. Factual and episodic knowledge was rated as being very similar. Explicit task ordering, - while informative is subjective and may not faithfully reflect cognitive mechanisms. We additionally analyse correlated patterns of responses and reaction times – a more implicit measure of similarity.

Clustering and Correlations

To gain convergence with explicit task ordering, we employed pattern similarity measures to interrogate the correlation patterns across two response measures - reaction times, and ratings. If tasks that originate from common or related cognitive processes, patterns of reaction times and response ratings could be expected to be similarly related.

Reaction times

Reaction times have been used to compare cognitive processing. Semantic priming paradigms demonstrate that reaction times are faster when a target word is preceded by a semantically related prime (Tipper, 1984). The time it takes to decide if two objects are

different, has been used as a measure object similarity (Podgorny & Garner, 1979). In this experiment, we evaluated whether *patterns* of reaction times are indicative of cognitive similarity. If a pair of tasks are produced by the related neural substrates, it would be reasonable to assume a relation in reaction times as well. Despite this reasonable hypothesis, we observed no relationship between cognitive task similarity and reaction times. One possible explanation of this is uncontrolled task difficulty. The measures we employed were somewhat robust towards differences by constant in reactions times. Since we investigated patterns of correlated RTs, if two cognitions are related, but differ by a constant time (of some milliseconds) we should still expect to see a correlation. However, our observed results show that patterns of reaction did not indicate task similarity. Additionally, they did not correlate with either patterns of responses or subjective task similarity. Taken together these results suggest that task similarity is not independent of the time it takes to provide a response, as measured by reaction times.

Response ratings

We demonstrated that patterns of response ratings can be a useful measure of conceptual similarity. Tasks that we sampled from the same domain of person-knowledge tended to covary across person-knowledge ratings. Additionally, they correlated with subjective similarity – providing convergence between explicit and implicit measures. Patterns of responses were not correlated with patterns of reaction times, suggesting that this measure captured task similarity that is independent from the time it takes to make a judgement. Our proof-of-concept analysis of people clustering demonstrated how people

are rated across person-knowledge domains. Across 40 people contained in the stimulus set, social and physical rating can be seen to covary. People rated less friendly, tend to be rated less trustworthy and less attractive.

Behavioral to Brain

In this chapter we demonstrate the feasibility to uncover the implicit structure of cognitive processing via multivariate analysis of behavioral patterns. Convergent results from subjective ordering and correlated patterns of responses reveal commonalities between different person-knowledge domains. Similarity between cognitions is independent of reaction times and may reflect common neural processes. Next chapter will address whether these cognitions share common neural substrates.

Conclusion

In summary, similarity analysis techniques are a promising approach for investigating cognitive processing. By looking at patterns of responses we are able to uncover similarities in the underlying processes. MIA and response ratings provide more insight into knowledge organisation than reaction time measures. By looking at behavioral response patterns we can plausibly recover implicit cognitive structure. We confirm a similarity structure between cognitive domains using two independent measure that yielded convergent results. This suggests graded commonality in neural substrates which we address in the next chapter.

Chapter 3. Regional Specialisation and Coordination within the Network for Perceiving and Knowing about Others

Introduction

In the preceding chapter we investigated the behavioural indicators of person knowledge. We observed that social and physical judgements are rated as being similar to each other and different from nominal, episodic and biographical knowledge retrieval. We also saw that this structure is reflected in the patterns of behavioural responses. Here we investigate the neural mechanisms when participants performed the same tasks. We will investigate whether tasks that were rated as being similar in the previous chapter rely on similar neural substrates.

Neuroimaging studies have identified an interconnected network of regions activated when we see and think about other people (Haxby, Hoffman, and Gobbini 2000; Gobbini and Haxby 2007; Fairhall and Ishai 2007). This network is composed of a perceptual 'core system', the occipital and fusiform face areas (OFA, FFA; Kanwisher et al. 1997; Gauthier et al. 2000) and the posterior superior temporal sulcus (pSTS). This core is complemented by an 'extended system' - a more loosely grouped set of regions implicated in a broad range of person-related cognition less related to perception (Haxby, Hoffman, and Gobbini 2000). The extended system includes lateral frontal regions: inferior frontal gyri (IFG), lateral orbitofrontal cortex (OFC); regions associated with internalised cognition: the anterior temporal lobes (ATL), precuneus, ventromedial and dorsomedial prefrontal cortices (vmPFC, dmPFC); and medial temporal regions: the

amygdalae and potentially the recently identified anterior temporal face patch AAFP (Moeller et al. 2008; Rajimehr et al. 2009).

Activation of the extended system has been linked to the spontaneous retrieval of person related knowledge (Todorov, Gobbini, Evans & Haxby, 2007). Seeing a familiar person brings to awareness a variety of related attributes: biographical facts (semantic knowledge), personal experiences (episodic knowledge), perceptual attributes (physical knowledge), trustworthiness (social knowledge) and their name (nominal knowledge). We use this wealth of person specific information frequently in our day-to-day lives. The question of how this diverse information is represented in the brain is an area of active research.

Numerous investigations of the contribution of these regions have produced diverse, and at times discordant, results (Table 1). For example, while the pSTS is classically characterised as a core 'perceptual' region, it is frequently linked with social and other person-related cognition that is unrelated to perception (Adolphs, 2003; Fairhall and Caramazza 2013a; Koster-Hale et al. 2017), potentially linked to the marked heterogeneity of the broader temporoparietal junction (TPJ, Gobbini and Haxby 2007; Hein and Knight 2008). Likewise, the varied functional attributions of the ATL to proper naming (Grabowski et al. 2001), semantic information (Tippett et al. 2000; Glosser et al. 2003; Olson et al. 2007) and social cognition (Frith, 2007) illustrate further the diversity and complexity of regional roles within the extended system (see Simmons and Martin (2009) for discussion).

Much insight has been gained from research addressing single cognitive function in one or a few brain areas. However, this approach has the potential to ambiguously define the true functions of regions. Uncertainties surrounding regional function are compounded by the tendency of research to address single cognitive functions in one or a few brain areas. Most regions of the extended system are recruited to some extent across multiple experimental contexts, including simple tasks such as repetition detection (Todorov et al. 2007) or superordinate categorization (Fairhall et al. 2014). In other words, most-all person related cognition involves most-all of the extended system. The important information may not be whether a specific region is modulated by a task but the *relative* change of regional activity across the entire network. Rather than attributing a function to a region or a region to a function it may be that the representation of cognition within the brain is best described in terms of the patterns of activation over distributed cortical networks.

The importance of network over region is especially important considering that many elements of the extended system fall within one prominent network in the brain, the internalised cognition brain network - a collection of brain regions associated with a range of internalised cognitive processes (Spreng et al. 2010). These include not only the eponymous 'default mode', task-deactivated states (Raichle et al. 2001) but also a broad range of internalised cognitive processes: general semantic knowledge (Binder et al. 2009; Fairhall and Caramazza 2013b; Huth et al. 2016); social cognition, (Greene et al. 2001; Van Overwalle 2009; 2011), as well as context integration episodic memory and mental time travel (Schacter and Addis 2007; Keidel, Oedekoven, Tut & Bird, 2017). As

elements of the internalised cognition network are frequently activated together, establishing specific contributions of each region has proved challenging (Moran et al. 2011; Van Overwalle 2011).

In this work, we isolate the network activated when we view familiar faces during a simple stimulus repetition detection task. Then we push the network towards different aspects of person knowledge (social, semantic, episodic, nominal and physical) to understand the relative importance of these processes to each region. The goals of the current experiment are twofold. The first is to re-examine the roles of key brain regions in cognitive function by considering these regions both in context of different cognitions and other brain regions. Specifically, we investigate the role of the anterior temporal lobe in social cognition and nominal knowledge and the role of the pSTS and AG in access to person knowledge. Our second goal is to apply network level representational similarity analysis (netRSA). NetRSA entails a multivariate approach based on the regional response magnitude within network nodes, with the multivariate element coming from the changing regional biases across different tasks. We employ netRSA to address two questions: a) how are cognitive domains represented across the network and b) how do these regions work together to accomplish the diverse range of ensemble functions of the network. We hypothesise that all person-knowledge regions are involved across all cognitive dimensions related to other people and that the cognitive flexibility of this system lies within subtle differences in the pattern of activation across the network.

Table 1. Illustrates complex cognitive landscape of the attribution of cognitive roles to various regions of the person-knowledge network. We sampled published reviews and meta-analyses reporting regional function in in the context of person perception of knowledge. Regional function attributions made by the authors of meta-analyses and reviews have been broadly grouped into the categories such as Perception, Semantic knowledge, or Working memory.

Brain region	Cognitive function
OFA	Perception (Adolphs 2002; 2003; Gobbini and Haxby 2007; Wieser and Brosch 2012)
FFA	Perception (Adolphs 2002; 2003; Gobbini and Haxby 2007; Minnebusch and Daum 2009; Wieser and Brosch 2012; Werner et al. 2013)
OFC	Motivation and reward (Bortolon et al. 2015) Top-down modulation of OFA, FFA (Bortolon et al. 2015) Emotion (Adolphs 2002)
IFG	Semantic knowledge (Gobbini and Haxby 2007) Working memory (Adolphs 2003) Perception (Adolphs 2003; Calder and Young 2005; Minnebusch and Daum 2009) Naming (Adolphs 2002) Attractiveness & Physical knowledge (Bzdok et al. 2010)
Precuneus	Familiarity (Adolphs 2003; Werner et al. 2013) Episodic knowledge (Adolphs 2003; Gobbini and Haxby 2007; Patterson et al. 2007; Wieser and Brosch 2012)
vmPFC	Social knowledge & personality traits (Adolphs 2003; Adolphs, 2002; Werner et al. 2013) Mental states (Gobbini and Haxby 2007; Wieser and Brosch 2012) Emotion (Werner et al. 2013)

	Attractiveness & Physical knowledge (Bzdok et al. 2010)
	Semantic knowledge (Adolphs 2003; Gobbini and Haxby 2007; Olson et al. 2007; Wieser and Brosch 2012)
	Identity representation (Adolphs 2003)
ATL	Naming (Calder and Young 2005; Olson et al. 2007; Patterson et al. 2007)
	Familiarity (Olson et al. 2007; Werner et al. 2013)
	Social knowledge (Olson et al. 2013)
	Perception (Adolphs 2003)
pSTS	Social knowledge & Mental states (Gobbini and Haxby 2007; Wieser and Brosch 2012)
	Emotion (Gobbini and Haxby 2007)
Amygdala	Physical knowledge (Bzdok et al. 2010)
	Social knowledge (Bzdok et al. 2010)
	Familiarity (Ross and Olson 2012; Heide et al. 2013)
ATFP	Identity representation (Ross and Olson 2012)

Methods

Participants

25 right-handed, native Italian participants (8 males; mean age: 23.2 years, range: 19-32 years) took part in this study. Participants had normal or corrected-to-normal vision and no history of neurological incidents. The study was approved by the University of Trento

ethical committee. All participants gave informed consent and were compensated for their time. Data from five subjects was excluded from further analyses, due to excessive motion during fMRI acquisition.

Stimuli

Stimuli were 40 pictures of famous faces and 40 pictures of famous buildings. The stimulus set consisted of Italian and foreign politicians, actors, singers and sportsmen, as well as landmarks (Eiffel Tower, Colosseum). Stimuli were cropped with a face-shaped mask and the eyes and mouth were aligned across faces. Stimuli extended 400 pixels vertically and 300 pixels horizontally and were presented centred on the screen (1280x1024 resolution, 60hz refresh rate), with grey background. After the experiment 13/20 participants were presented with faces, they saw in the experiment and asked whether they recognised the celebrity. On average, the subjects recognised $M = 84\%$ of faces.

Task

Each experimental block started with 4s instruction screen specifying the task, followed by 6s of fixation cross. After that a face was presented for .5s followed by 2s of fixation cross during which subjects provided a response via button box. Within each 8-trial block, participants were instructed to respond to questions covering five categories of person knowledge: episodic memories, semantic knowledge, social judgments, nominal knowledge and physical knowledge. For each of the categories, we chose two different

probe questions that require access to each kind of knowledge (totalling ten experimental tasks; see Figure 2.1 and Table 2). In addition, there were two 1-back matching control tasks with either pictures of faces or famous monuments. The experiment consisted of five runs (8 min, 42s each). Sixteen blocks were presented in a randomised order (one block for each task plus three face and three monument 1-back control blocks).

Table 2: Experimental questions. We selected five categories of person related knowledge (nominal, perceptual, episodic, social and semantic). For each category we chose two different probe question that require accessing the specific category of knowledge. Each task was presented in a block of eight trials. On each trial, participants were instructed to relate the task question to the famous person.

Knowledge category	Task Cue	Participant instructions	Answer choices
Nominal	Common name	How common is this person's name?	Likert scale (1-4)
	Full name	How well can you recall the person's full name?	Likert scale (1-4)
Perceptual	Attractive	How attractive do you find this person?	Likert scale (1-4)
	Distinctive	How distinctive is this person's face?	Likert scale (1-4)
Social	Friendly	How friendly is this person?	Likert scale (1-4)

	Trustworthy	How trustworthy is this person?	Likert scale (1-4)
Episodic	Familiar	How familiar is this person is to you?	Likert scale (1-4)
	First memory	For how long have you known this person?	Likert scale (1-4)
Semantic	How many facts	How many facts could you recall about this person?	Likert scale (1-4)
	Occupation	What is this person occupation?	Predefined categories (see 'task')

Participants answered 9/10 questions using a 1-4 Likert scale. Occupation question (“what is this persons’ occupation”) had predefined categories (1 = actor or TV presenter, 2 = singer or musician, 3 = politician or sportsman, 4 = none of the above). Prior to scanning, participants practiced answering experimental questions on a different set of famous people repeating each question for five trials.

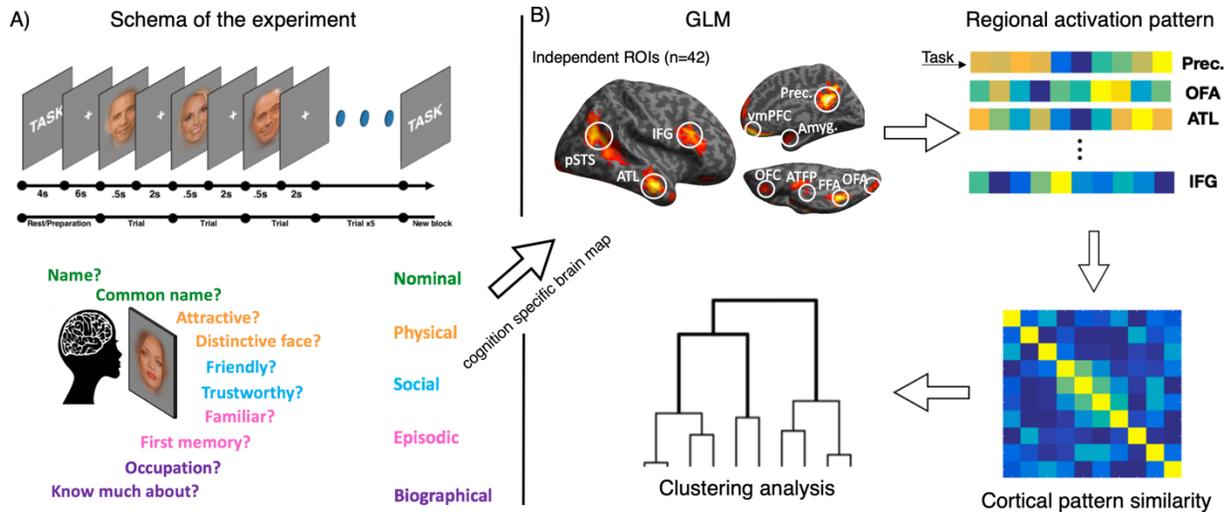


Figure 2.1. (A) Top: Schematic representation of the experiment. Experimental blocks were preceded by 4s of instruction screen, and 6s fixation point. Each trial consisted of .5s face presentation and 2s fixation. Bottom: five domains of person knowledge and two questions per domain are noted in corresponding colours. (B) Data analysis schematic. ROI beta averages for each of the ten tasks were extracted from ROIs, correlated and subjected to RSA.

Data Acquisition

Participants were scanned at the Center for Mind/Brain Sciences (CIMeC), University of Trento, Italy. Data was collected using Bruker BioSpin MedSpec 4T, with 8-channel phased-array head coil. Five runs of 209 echo-planar volumes, consisting of 34, AC-PC aligned axial slices were acquired while participants performed the task (FOV = 64mm x 64mm, TR = 2.5s, TE = 33, FA = 73°). Voxel size was 3x3x3mm with a 1mm gap. In addition to functional data, a whole brain T1 MPRAGE anatomical image was acquired (whole brain (FOV = 256x224, 176 1mm axial slices).

Region of interest definition

Regions of interest (ROIs) were selected from an independent (N=42) experiment, conducted for high power functional localisation. In the localiser experiment participants performed a 1-back matching task with 12 second blocks of famous faces, common animals or common objects. The contrast faces > animals+tools ($p < .05$ FWE corrected) was used to identify face selective peaks (Table 2). 7.5mm radius spheres were drawn around the peak voxels and task evoked brain responses (beta estimates) were extracted for each subject.

To investigate differences in pSTS/Angular gyrus (see: introduction) we anatomically constrained the volume with angular and superior temporal gyri masks. Then the masks were inflated, and the overlap was removed. This allowed us to distinguish face selective anatomical activation within anatomical angular and superior temporal gyri within anatomical boundaries.

Monuments > Face control tasks contrast was used to validate ROI selection. We expect extended system, but not core systems ROIs to respond to person-knowledge tasks. To ensure that core system ROIs are correctly localised we will compare face selective response (face > monuments control tasks) in the core system (see below: Accessing person related knowledge recruits extended, not core, components.)

Table 3: ROI sphere centre coordinates. Peak coordinates for regions active in the localiser experiment (N=42) and ROI sizes in voxels after thresholding. Coordinates are in MNI space.

Region	Hemisphere	X	Y	Z	ROI Size (voxels)
Precuneus	Medial	3	-52	29	81
OFA	Right	30	-91	-10	65
	Left	-33	-88	-10	49
FFA	Right	42	-46	-22	81
	Left	-39	-46	-22	30
IFG	Right	39	17	23	44
	Left	-36	20	26	38
ATL	Left	-60	-7	-19	69
	Right	57	-7	-19	81
Amygdala	Left	-21	-10	-13	62
	Right	21	-7	-16	59
dmPFC	Medial	6	59	23	59
vmPFC	Medial	3	50	-19	66
OFC	Right	33	35	-13	58
	Left	-33	35	-13	27
ATFP	Right	33	-10	-40	39
	Left	-36	-10	-34	24
Angular	Left	-48	-67	35	68
	Right	42	-64	35	57
pSTS	Left	-48	-49	14	54
	Right	48	-55	14	71

Data Analysis

Data were pre-processed with SPM12. Functional images were realigned to account for motion, grey matter segmented, warped into common space and smoothed with 8mm

FWHM kernel. Subject specific response estimates (beta weights) were derived by fitting a general linear model (GLM) to the data. 12 regressors (10 tasks, 2 controls) were included as explanatory variables. Six motion parameters from re-alignment procedure were included as regressors of no interest. We drew 7.5mm spheres around chosen coordinates (Table 2) and extracted the mean beta value significantly active at $p < 0.001$ within those ROIs (contrast faces > animals+tools). To isolate the magnitude of cognitive response, we subtracted beta value for 1-back matching face control task from each experimental task.

Multivariate Analyses

ROI responses across ten tasks were averaged across voxels and correlated to obtain a dissimilarity matrix ($1-r$), which was then subjected to Ward hierarchical agglomerative clustering. For task similarity analysis the matrix was transposed before correlating so that similarity matrix consisted of task correlation across 21 ROIs. Similarity models were constructed and fitted according to conventional RSA framework (Kriegeskorte, Mur & Bandettini, 2008) by building a matrix of expected similarity (either regional similarity or task similarity) with 1 denoting expected pairwise similarity and 0 denoting no-relationship. To establish model fit we correlated the model with observed network RSM (r) for each subject. Subject-model correlations were then subjected to one sample t-test (null hypothesis $r=0$) to establish significance.

Results

Behavioural Data

Mean Reaction time (RT) was $M = 1203\text{msec}$, $SD = 111\text{msec}$. Subjects reacted fastest during full name task ($M=1110\text{ms}$, $SD=120\text{ms}$), and slowest during common name task ($M=1386\text{ms}$, $SD = 200\text{ms}$). RTs differed across the ten tasks ($F(9, 171) = 17.13$, $p < .001$). Critically, this RT effect did not persist when tasks were collapsed into the five domains of knowledge used in the imaging analysis (i.e. “Nominal“, “Physical“, “Social“, “Episodic“, “Semantic“; $F(4, 95) = 2.21$ $p = .076$). RT did not significantly differ between face ($M = .71$, $SD = .11$) and place ($M = .68$, $SD = .12$) control 1-back matching tasks ($t(18) = 1.96$, $p = .066$).

Mean ratings ranged from $M = 2.2$, $SD = 0.25$ (attractiveness task) to $M = 3.1$, $SD = .38$ (full name task). Ratings differed across 10 exemplar tasks ($F(4.8, 91.7) = 27.741$, $p < .001$, Greenhouse-Geisser correction) and 5 cognitive domains ($F(4,76) = 20.03$, $p < .001$). Faces were rated highest on attractiveness ($M = 2.79$, $SD = .025$) and lowest on participants' ability to recall full name ($M = 1.82$, $SD = .38$).

Accessing person related knowledge recruits extended, not core, components.

To assess the global importance of access to person knowledge in the core and extended systems for person perception/knowledge, we compared the average regional increase when participants accessed the 10 variants of person knowledge, compared to the one-back matching task on famous faces. The results are shown in Figure 2.2.

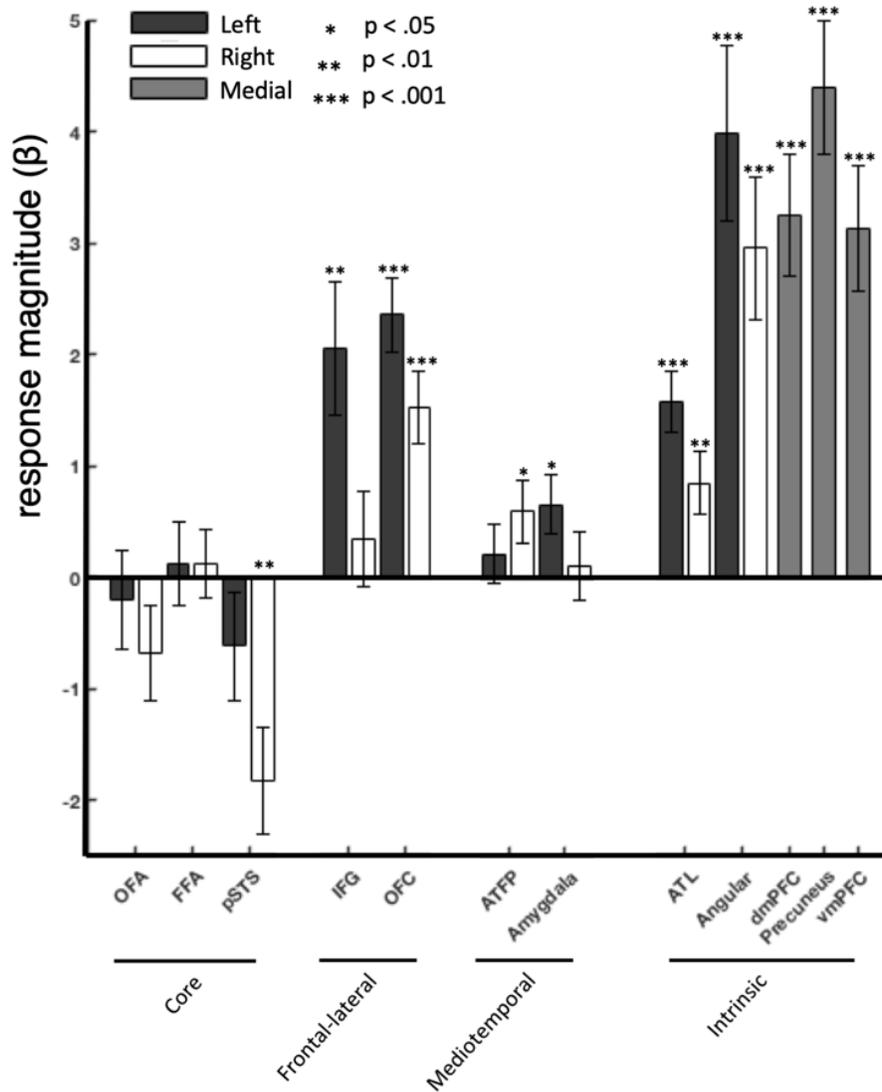


Figure 2.2: Regional response to person knowledge access (average of experimental tasks > face control task). Bars show response estimate magnitude (beta value), error bars represent standard error of the mean (SE). Retrieving person knowledge activated extended system regions to variable extent, but core system did not show an increased response.

The most apparent distinction is between regions of the core system, which show no increase in activity ($t(19) = -1.69, p = .11$; averaged across all core regions) and other

elements of the network ($t(19) = 6.98, p < .001$; averaged across all extended regions). While unresponsive to person-knowledge tasks, core system responded more to face than monument control tasks, validating ROI selection, $t(19) = 2.91, p = 0.009$.

A second clear organisational feature of person knowledge is the left lateralisation of this process, with left hemisphere regions showing a greater relative increase when accessing person knowledge than their right hemisphere counterparts ($t(19) = 4.67, p < .001$).

Individually, regions of the internalised cognition part of extended system were all strongly recruited during access to person knowledge ($t\text{-values} > 4.60, p < .001$). Lateral frontal regions, the IFG and lateral OFC show an increased response most consistently in the left, with the right IFG failing to show a significant modulation. More subtle increases are seen in medial temporal lobe structures, with only the left amygdala and the right ATL exhibiting significant modulation when person knowledge was accessed.

The Role of Cognitive Domain across the Person Knowledge Network

To investigate the role of different kinds of knowledge in the person knowledge network, Figure 2.3 shows the scaled importance of each cognitive domain in those regions that were activated during access to person knowledge. Here we focus only on regions that showed at least a significant unilateral increase when person knowledge is accessed. For simplicity, we have collapsed across hemisphere (significant interaction between domain and hemisphere were present only in IFG; $F(4,76) = 2.508, p = 0.049$). Regions of the

intrinsic network are strongly involved in all cognitive domains with the exception of nominal knowledge, to which ATL and dmPFC were unresponsive.

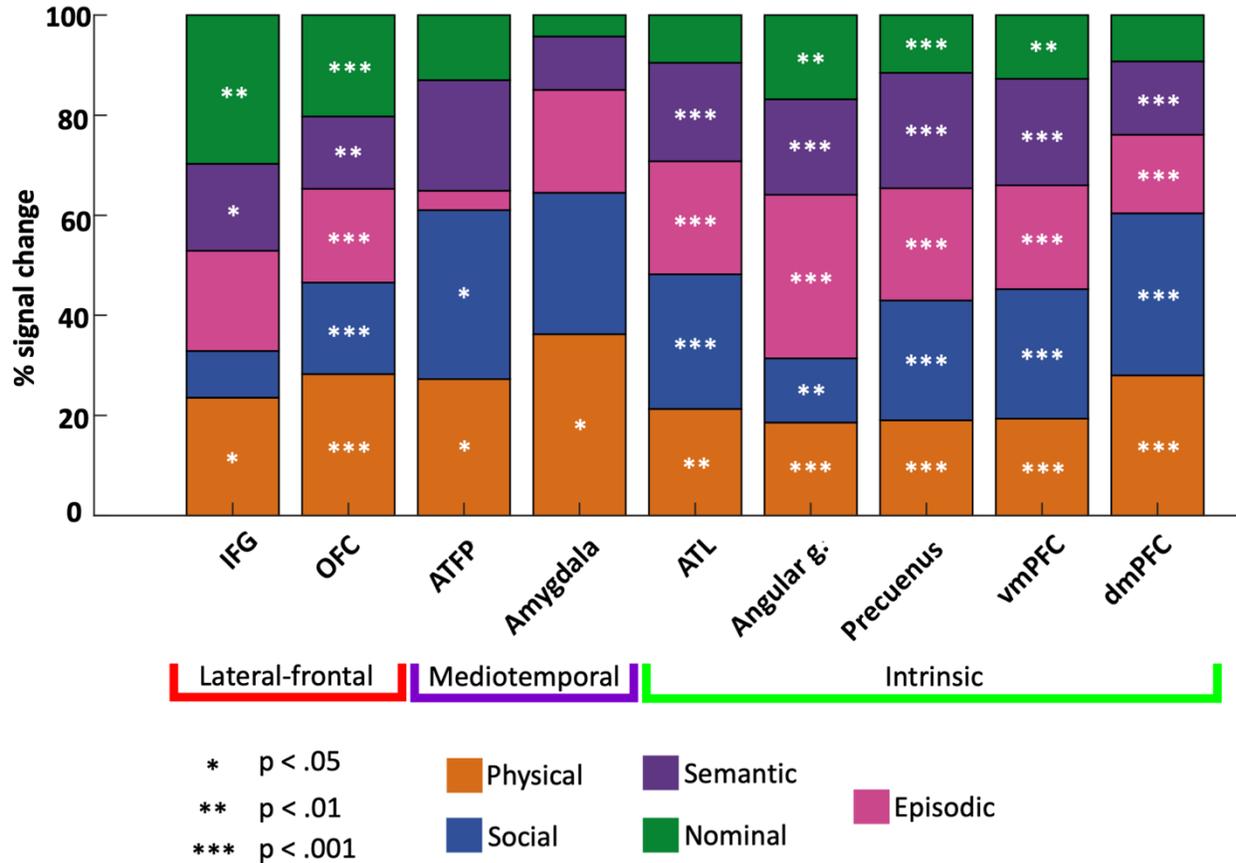


Figure 2.3. Regional preference patterns. Percentage of total activation elicited by each cognitive domain. Stars denote significance threshold. Most regions are involved in most cognitive domains. Although patterns vary across components of the network, it can also be seen that regions that respond to episodic knowledge tend to respond to semantic access as well.

The ATL is weighted towards Social not Nominal knowledge.

The involvement of ATL in social or nominal (proper names) knowledge is a matter of current contention (Grabowski et al. 2001; Olson et al. 2013). We exploited the presence of all these cognitive aspects within one single study to address their relative importance in bilateral ATL. ANOVA revealed significant differences across cognitive domains ($F(4, 95) = 8.13, p < .001$). A planned comparison specifically testing the role of social and nominal tasks revealed that access to social knowledge recruited this region to a greater extent than access to nominal knowledge ($t(19) = 5.79, p < .001$), with nominal knowledge showing marginal increase compared to control ($t(19) = 2.08, p = .052$). Follow up analyses revealed that, while the largest response was evident during access to social knowledge, this was not significantly greater than episodic, semantic or physical cognitive domains, each of which showed a greater response than the nominal tasks (post hoc: all t -values > 2.49 , all $p < .02$, uncorrected). The response during the nominal tasks were not diminished in all regions, with IFG responding more strongly to access to nominal than social knowledge ($t(19) = 3.1, p = .006$). These results show a distinction between the way social and nominal knowledge is processed, social tasks engage ATL more than nominal ones, while the IFG exhibits the opposite pattern, being more responsive to nominal than social knowledge.

Perceptual and non-perceptual processing in the pSTS/Angular gyrus

Here we sought to test whether the pSTS and AG, have diverse functional roles. While these regions sometimes form a contiguous activation cluster, pSTS has been implicated in perceptual processes, while AG is involved in cognitive processes such as intent and

mental state processing (Gobbini & Haxby, 2007). The functional division between these two regions was apparent even at a level of global access to person-knowledge, (c.f. figure 2.2). Specifically, while the pSTS showed a suppressed response during access to stored person-knowledge (pSTS left: $t(19) = -1.26$, *ns.*; right: $t(19) = -3.81$, $p = .001$), the angular gyrus conversely showed a robust response (AG left: $t(19) = 5.81$, $p < .001$.; right: $t(19) = 2.99$, $p = .008$). This difference was also apparent in the global laterality pattern. pSTS showed a stronger inhibition in response on the right ($t(19) = 3.47$, $p = .001$), consistent with the pattern of right laterality in core regions. Conversely, there was no laterality effect in global response of the angular gyrus ($t(19) = 1.57$, $p = .134$). Collectively, these patterns underscore the pronounced functional subdivision of these regions of the temporo-parietal junction.

At the level of the single cognitive domain (see Figure 2.3) an increased response in AG was evident across each cognitive domain (t -values > 3.12 , $p < .006$) while the pSTS did not show an increase for any cognitive domain. Finally, while cognitive domain did not have a variable effect in pSTS at the univariate level, the angular gyrus responded more during episodic memory retrieval than any other task (t -values > 2.8 , $p < .01$). Together, these results demonstrate a pattern of response in the pSTS consistent with core perceptual processing and a pattern in the AG consistent with access to knowledge. Differences in inter-regional coordination patterns will be discussed in the next section.

Inter-Regional Coordination and Network organisation

Which brain regions work together to accomplish the person-knowledge network's varied functions? The functional coordination between ROIs was examined through a network level Representational Similarity Analysis (*netRSA*). The ten task-induced beta patterns were correlated between each pair of ROIs and subjected to hierarchical clustering analysis (see figure 2.1 and methods). The validity of this *netRSA* approach is confirmed by the close proximity of left and right regional homologues (figure 2.4): In all cases, despite the anatomical distance, a given ROI's cognitive profile matched most closely to that of its contralateral counterpart. An RSA template model (figure 2.4), confirmed the high reliability of this effect across subjects ($t(19) = 13.89, p < 0.001$). This result highlights the commonality of function between hemispheric homologues despite hemispheric asymmetries in the overall response.

At a descriptive level, *netRSA* revealed the expected cognitive clustering of core regions (OFA, FFA, pSTS). Interestingly, the lateral frontal regions the orbito-frontal and inferior frontal gyri cluster with the core system, rather than the other components of the extended system (figure 2.4 – red cluster). To test whether lateral frontal regions coordinate more closely with core or extended systems, we built competing models of regional coordination (see figure 2.4). Post hoc comparisons of the goodness of fit between the two competing models confirmed that lateral frontal regions co-ordinate more closely with core than other extended system regions ($t(19) = 3.09, p = .006$). Considering regions separately, this effect persisted for IFG ($t(19) = 4.35, p < 0.001$), while no preference was evident for OFC ($t(19) = 1.59, p = 0.13$).

Within the extended system, regions overlapping with those associated with internalised cognition form a distinct cluster (Cluster 2 - green) with respect to the amygdala and anterior face patch of the medial temporal lobe (Cluster 3 – purple). However, planned comparison of whether the ATFP groups more closely with the core or extended system resulted in no evidence for either hypothesis ($t(19) = -0.21, p = 0.84$).

Network organisation: regional pattern similarity

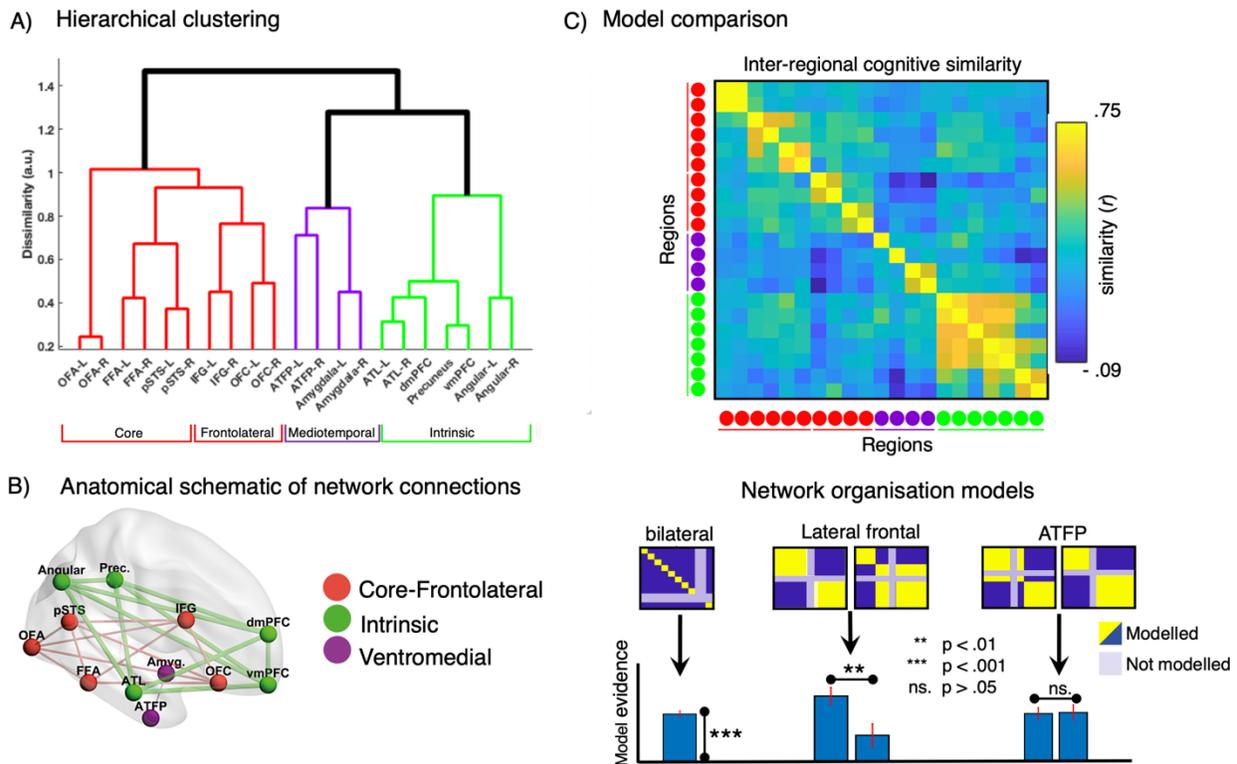


Figure 2.4. A) ROI clustering. First major division separates core together with IFG & OFC from the rest of the extended system (red vs (green + purple)). Within the extended system, a further division is evident between medial temporal regions and regions involved in the intrinsic system (purple, green). B) Schematic representation of clusters projected onto the brain. C) model comparison schematic. Competing models of network

organisation were constructed, fitted to observed data for each participant, and compared in a paired samples t-test.

Cortical Similarity between Cognitive Domains

To address the fundamental question of how different forms of cognition relate to one another, netRSA was performed across ROIs to investigate how similar the neural representations of cognitive domains are in the brain (see Figure 2.5 and methods). Despite variance across tasks, reactions times, and in one case response scales, the task pairs for each cognitive domain (e.g. ‘common name’ and ‘full name’ for nominal knowledge) are grouped together. This illustrates the efficacy of netRSA in this context, providing another internal validation. To test our a priori selection of task couplets, we built a model to test their similarity. Results show that tasks from the same cognitive domain have highly similar representation across the person knowledge network ($t(19) = 4.14, p < 0.001$).

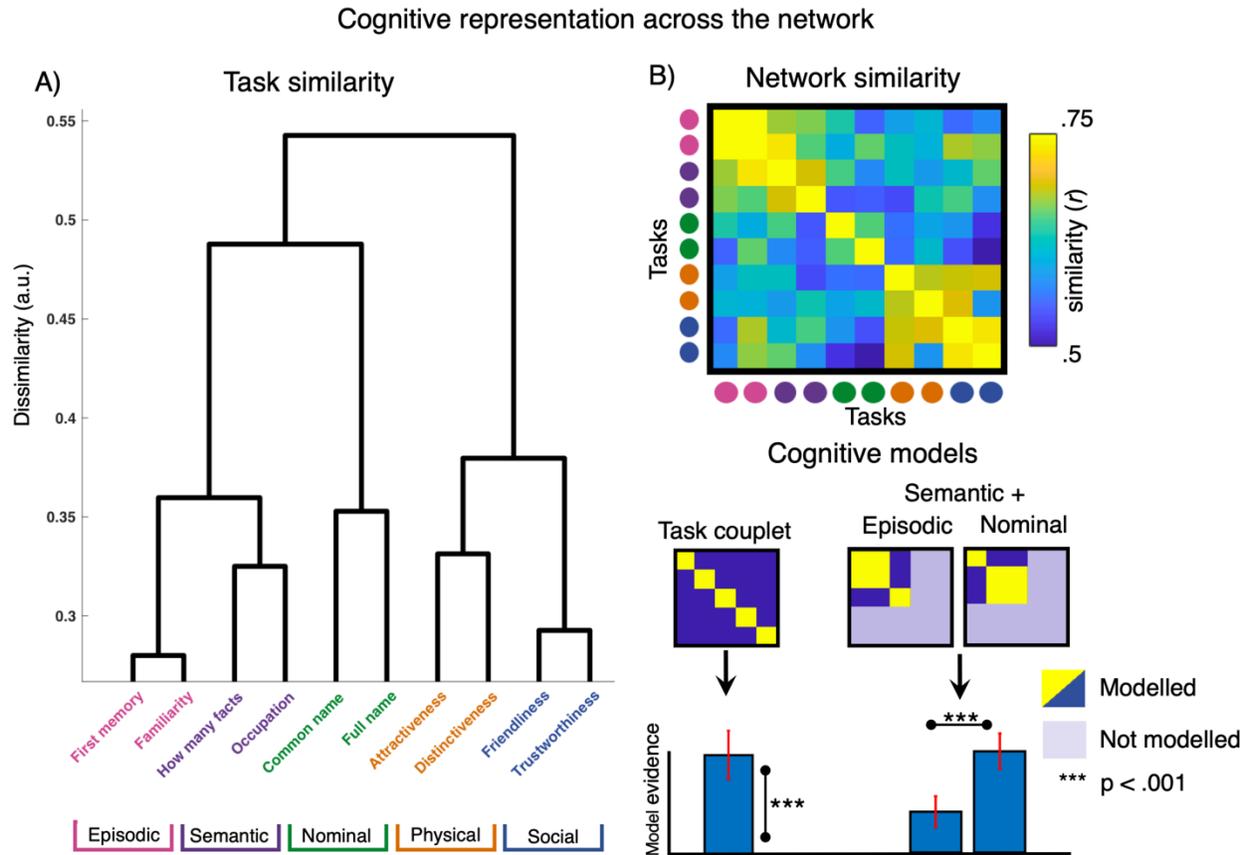


Figure 2.5. A) Task similarity in core and extended system ROIs. Tasks are grouped according to the domain they were sampled from. Episodic and semantic knowledge retrieval tasks elicit differentiable patterns from nominal, physical or social ones (dendrogram, left). B) fMRI pattern similarity matrix and models tested.

At the next level of hierarchical clustering, results reveal three distinct cognitive clusters. Network activity is more similar for physical (yellow) and social (blue) knowledge than to other forms of knowledge. Likewise, episodic (pink) and semantic (purple) knowledge form a cluster, that is distinct from nominal knowledge. We compared competing models (see Figure 2.5) to make inference about whether semantic knowledge more closely relates to the episodic or nominal domain. Results

confirm that semantic knowledge is more similar to episodic memories, than nominal knowledge ($t(19) = 3.97, p < .001$).

Whole Brain Analysis - Beyond the person-selective network

The motivation of this project was to gain insight into the normal function of the well-characterised network for person knowledge by strongly pushing the system towards access to different domains of knowledge. While we focus on the network for perceiving and knowing about others, cognitive processes are not enacted solely by the person knowledge system. It is important to consider that these systems presumably pair and couple with brain regions outside the person-selective network, with the network's periphery potentially driving transient specialisation within the network itself.

A whole brain analysis comparing each cognitive domain to the average of the other domains is presented in figure 2.6. Notably, no regions demonstrated a significant preference for two or more cognitive domains, consistent with non-overlapping cognitive specialisation outside the person knowledge network. Access to nominal knowledge is characterised by a broad pattern of activation, stronger in the left hemisphere than the right (Figure 2.6, Table 4). It encompasses left hemispheric sites associated with language production and comprehension (Price, 2012). Conversely, accessing physical knowledge is more pronounced in the right hemisphere, predominantly in the superior portion of the IFG. Social knowledge retrieval more strongly activated left ATL and an anterior patch of dmPFC. Recalling episodic memories involved the posterior cingulate cortex (PCC), angular gyrus as well as right lateral frontal pole and bilateral patches of

the superior frontal gyrus. Semantic knowledge retrieval tasks did not selectively recruit any region outside of the person-knowledge network.

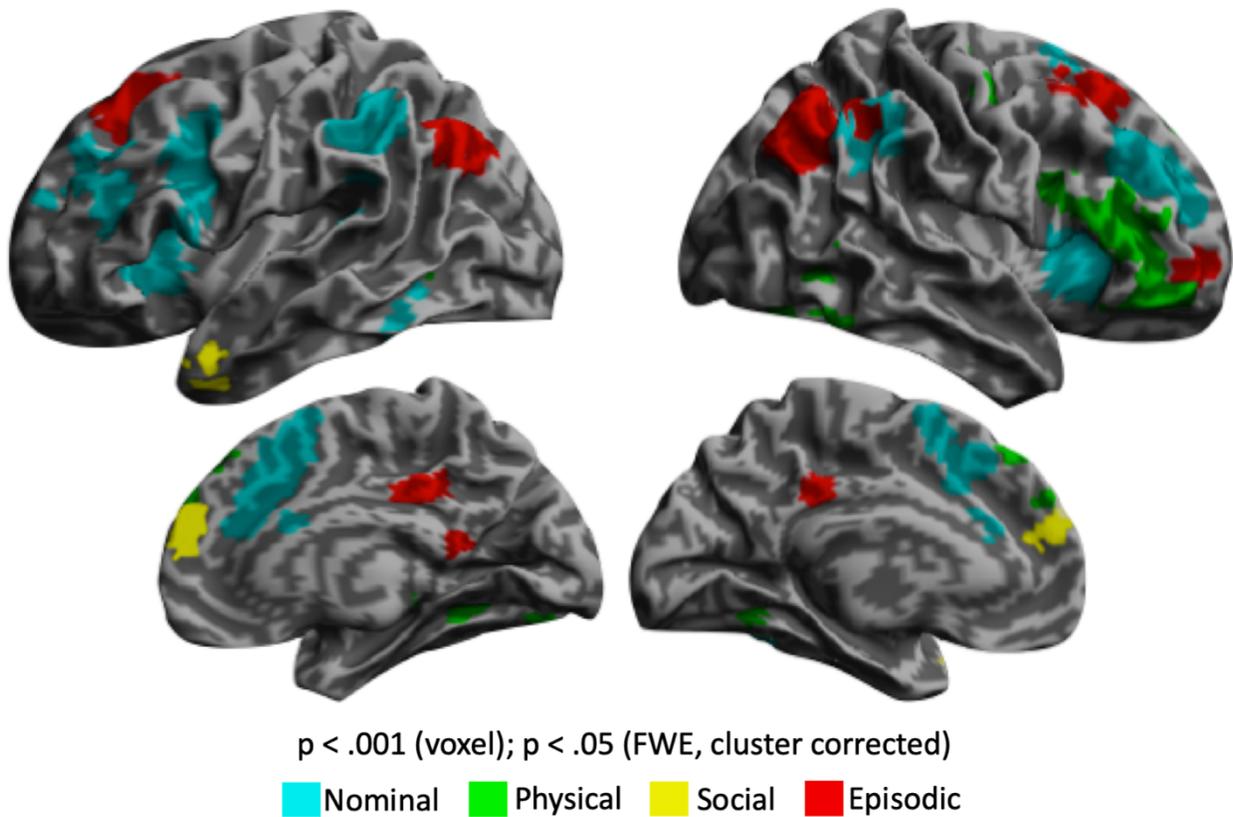


Figure 2.6. Cognition specific brain activations. Whole brain map highlights peripheral, domain specific cognitive systems that are recruited during diverse kinds of person-knowledge retrieval. Semantic knowledge did not elicit significant clusters of activity.

Table 4: Peak location, extent and cluster-level significance of whole brain analysis of the differential effects of each cognitive domain (see also: Figure 2.6). SFG - superior frontal gyrus, IPL - inferior parietal lobule. PMC - premotor cortex, PCC – posterior cingulate cortex.

Domain	Region	Hemisphere	cluster		Peak
			p (FWE)	size	T
Episodic	Angular	Right	< 0.001	507	8.06
	Angular	Left	< 0.001	264	6.57
	SFG	Right	< 0.001	216	5.68
	PCC	Medial	< 0.001	142	5.68
	SFG	Left	0.001	125	4.77
	OFC	Right	0.017	71	4.65
	Precuneus	Medial	0.032	61	4.16
Social	dmPFC	Medial	< 0.001	396	6.57
	ATL	Left	0.001	130	4.58
Physical	IFG	Right	< 0.001	616	5.72
	FFA	Right	< 0.001	274	4.86
	PMC	Right	0.006	87	4.38
	Amygdala	Right	0.020	68	4.37
	dmPFC	Medial	< 0.001	253	4.34
	FFA	Left	0.004	94	4.24
	IPS	Right	0.005	92	3.93
Nominal	IFG	Left	< 0.001	1188	6.99
	SFG	Medial	< 0.001	1670	6.40
	Angular	Left	< 0.001	370	6.35
	Angular	Right	< 0.001	265	5.31
	FFA	Left	0.003	97	4.70

Discussion

Cognitive processes are distributed across networks of regions. By having our participants perform a broad range of cognitive tasks and exploiting data complexity through derivations of statistical tools such as RSA, we could focus on the subtle differences in inter-regional coordination that endow distributed cortical networks their considerable cognitive flexibility. By relating task activity not only to a baseline but to other tasks, we were able to establish regional preferences which allowed us to tackle specific hypotheses about regional function. Leveraging multivariate methods provided insight into regional coordination, uncovering principles of network organisation, as well as the similarity between different cognitive processes allowing us to build a cognitive taxonomy grounded in brain representation.

Accessing person related knowledge recruits extended, not core, components.

Clear functional division between core and extended system can be seen even at the broadest view of systems organization. Face-selective core regions (OFA, FFA & pSTS) were not responsive to person knowledge tasks. This stands in contrast of the extended system regions which were strongly engaged by the experimental tasks, and particularly so in the left hemisphere. This is consistent with a predominant role of these regions in the extraction of perceptual information which is made available to the other elements of the system (c.f. Downing and Peelen 2011).

Functional subdivisions between pSTS and Angular Gyrus

The pSTS was originally designated as part of the core system (Haxby, Hoffman, and Gobbini 2000b). Over the years this classification has become less clear and pSTS has been reclassified as part of both core and extended systems (Gobbini and Haxby 2007). The pSTS is part of a heterogeneous cortical region, the temporal parietal junction (TPJ), which includes pSTS, the angular gyrus and the supramarginal gyrus. In this study, we anatomically divided the person selective patch into pSTS and angular gyral components. We observed a pronounced dissociation across this subdivision. Firstly, compared to the face repetition-detection control task, AG exhibited a strong global response to access to person related knowledge while pSTS was either unresponsive (left) or suppressed (right). Secondly, the response profile across cognitive tasks grouped them differently, with pSTS clustering with other core regions while AG clustered with regions of the extended system associated with intrinsic cognition.

These results indicate that the contiguous TPJ activation reported during the viewing of familiar faces is functionally heterogeneous, divided into an angular gyrus component that responds to non-visual theory of mind (Saxe and Powell 2006) and amodal access to person knowledge (Fairhall and Caramazza 2013a) and a pSTS component involved in face perception. Future work will determine how these anatomically adjacent brain regions coordinate across other cognitive processes but within the context of person-related cognition, they appear to be highly distinct.

Regional tuning of ATL

Different hypotheses propose that ATL might mediate person-specific social knowledge (Olson et al. 2007) or proper naming (Semenza 2011). Here we directly contrasted these two hypotheses and observed that ATL responds strongly to social knowledge and is unresponsive to nominal knowledge retrieval. This is in apparent contrast to early positron emission studies which show strong activation of the region during overt naming (Gorno-Tempini et al. 1998; Grabowski et al. 2001). However, it is consistent with the finding that the ATL responds equally to familiar people whether or not the name is known by the participant (Gesierich et al. 2011). The reason for these disparities may lie in fMRI signal drop-off in the ATL (Devlin et al. 2000) or in specific representational or phonological demands associated with overt speech production. Alternatively, it may be that ATL activation observed in earlier studies was not specific to nominal knowledge but rather, to generalised activation of person knowledge. Indeed, this generality is evident in the present study where, access to semantic, episodic and physical cognitive domains, as well as social, activated ATL. It is noteworthy that, while social knowledge did not produce a stronger response in the ATL ROI than these other three domains, the whole brain analysis (Figure 2.6) only social knowledge showed a significant cognitive selective response in adjacent the left ATL, consistent with the importance of social conceptual knowledge in this region (Simmons & Martin, 2009).

Regional coordination across the person knowledge network

Elements of the network for person knowledge coordinate to form its diverse range of functions. Here, in contrast to investigating functional connectivity over time (e.g. Fairhall and Ishai 2007) we consider how these regions functionally coordinate over different tasks in response to their varied cognitive domains. Consistent with classic models (Haxby, Hoffman, and Gobbini 2000b) we observed that functional coordination between core perceptual (OFA, FFA and pSTS) regions was high. Interestingly, we observed that the fronto-lateral components of the classic extended system, IFG and OFC, coordinated more closely with these core regions rather than other elements of the classic extended system. It is notable that these fronto-lateral regions, particularly IFG, are closely related to extrinsic, task-activated, networks, distinguishing them from extended regions associated with the anti-correlated intrinsic resting state network (Fox et al. 2005). Additionally, during access to stored knowledge, the fronto-lateral IFG is implicated in guiding access to relevant information (Martin and Chao 2001; Wagner et al. 2001; Thompson-Schill 2003) suggesting a modulatory rather than representational role.

Other components of the extended system appear to coordinate most closely with each other across different cognition domains. Hierarchical clustering revealed an apparent dissociation between medial temporal components (ATFP, amygdala) and those associated with internalised cognition (vmPFC, precuneus, ATL, AG). However, this was not confirmed by statistical analysis and future work will be needed to verify this grouping. A planned comparison of whether ATFP grouped more closely with core or extended systems revealed no evidence in either direction.

Cognitive taxonomy in person knowledge

By considering the similarity between the neural profile of different cognitive domains we can gain insight into the relationship between these cognitive processes. We implemented this approach within the network for perceiving and knowing about others by comparing the profile of each task across the 21 ROIs comprising this network. We observed that for each task, despite variations in reaction times and task structure, that the two task-exemplars for each of the five cognitive domains reliably clustered with their counterpart. Demonstrating that cognitive domain is the primary grouping factor of activation patterns across this network and validating our selection of tasks. The general pattern of cognitive clustering across the regions suggests that social and perceptual knowledge share similar neural patterns, as do semantic and episodic knowledge, with nominal being represented somewhat distinctly. These findings fall broadly within hypothesised domain-specificity boundaries (Spunt and Adolphs 2017) suggesting that declarative memory (episodic, semantic tasks) and language (nominal tasks) are part of the 'cognitive' macro-domain, while facial reception (physical tasks) and theory of mind (social tasks) are part of the 'social' macro-domain.

Of specific interest to us was the relationship between episodic, semantic and nominal knowledge. One classic distinction in forms of declarative memory is between episodic (personal experience) and semantic (general knowledge) (Tulving 1972). The term 'semantic knowledge' refers to a broad range of knowledge about objects, factual knowledge and linguistic access to word meaning. Here we specifically contrasted

competing models comparing whether semantic memory factual knowledge about people was more closely related to our personal experience and memories about that person than to access to their name, a task domain that strongly recruited language circuitry (see next section). Model comparison revealed significantly greater evidence that semantic memory clustered more closely to episodic than nominal access. This grouping is roughly apparent in the tuning profiles presented in Figure 2.3, where episodic and semantic domains load most heavily on regions associated with intrinsic cognition while nominal knowledge follows a different pattern – additionally engaging regions outside of the intrinsic network (e.g. IFG). In a broad sense, this result suggests that within the context of the tasks used in this study, semantic access shares a mechanism with episodic rather than linguistic neural systems.

Cognitive domain across the whole brain

In the study, we sought to understand how different elements of the person knowledge network, active spontaneously when we view familiar people (Gobbini and Haxby 2007), contribute to our diverse array of person related knowledge. To this end, we perturbed the system towards five different cognitive domains to understand regional processing bias. However, these cognitive domains are not manifest solely within the person knowledge network and a whole brain analysis revealed the cognitive-domain selective regions outside this network. Figure 2.6 shows the selective activation of cognitive domains across the brain (domain v. others).

Recruitment of language regions in the Nominal tasks: the supramarginal gyrus, lateral PFC and dorsomedial PFC (Price 2012), validate the importance of linguistic processes in the performance of this task. Similarly social knowledge tasks preferentially recruited regions of the left ATL and an anterior section of dorsomedial PFC, consistent with the social cognition network (Adolphs 2009). The engagement of broad parts of right lateral PFC as well as small clusters in medial temporal and occipital areas is consistent with the involvement of these regions in the retrieval of perceptual attributes (Kan et al. 2010). Episodic access generally recruited regions associated with autobiographical memory (Schacter et al, 2012; Spreng & Grady, 2009) In particular, bilateral angular gyrus showed a strong modulation with episodic tasks and is an area associated with the strength of autobiographical recollection (Rissman et al., 2016). Notably, despite the involvement of a broad set of regions consistent with the episodic domain, the hippocampus was not recruited, even at uncorrected thresholds. The reason for this absence is uncertain but may relate to the use of a familiarity rather than recognition weighted tasks in our design (Eichenbaum, Yonelinas & Ranganath, 2007) or to other factors related to this kind of episodic access. Collectively, these results indicate the separable nature of the cortical processes associated with these five cognitive domains and suggest that cognitive biases within the elements of the person knowledge network may arise from coordination with brain regions outside the network.

Conclusion

The current study examined how the distributed cortical network for perceiving and knowing about others coordinates to accomplish its diverse range of cognitive functions. By examining a range of tasks within a single study we were able to observe a clear preference for access to social over nominal knowledge in ATL and, within the TPJ, a clear division between a more perceptual pSTS and a more cognitively weighted angular gyrus - both in terms of their global response to cognitive access and the profile of activation across different cognitive domains. Through further multivariate analysis of the cognitive response profile across the network, we observed that regional coordination profiles grouped core regions with the lateral frontal extended system regions, which functioned relatively independently of intrinsic components of the extended system. Moreover, we were able to use similarity in activation patterns across the network to reconstruct a taxonomy that reflected how similar cognitive domains are to one another within the cortex. Notably, we observed that access to factual semantic knowledge employs neural substrates more similar to episodic memory than to language-related nominal knowledge. Collectively, these results demonstrate the importance of network level dynamics in the instantiation of person related cognition and knowledge

Results of this chapter provide novel insight into both knowledge organisation and regional coordination. To rule out the perceptual effects on our findings, next chapter will investigate whether these findings are robust across stimulus modalities. If what we found is a true organisation of knowledge, we should see it regardless of stimulus modality.

Chapter 4: Organisation of person knowledge across stimulus modalities

Introduction

Previous chapter investigated how person knowledge is organised in the cortical network for face perception across core and extended systems. We described correlated organisation of knowledge and observed that episodic memories and biographical knowledge are represented distinctly from nominal knowledge. Similarly, Social and Physical knowledge retrieval relied on different neural substrates. Regionally, perceptual (OFA, FFA, pSTS) co-ordinate with frontal regions (OFC, IFG) across different kinds of person-knowledge retrieval. Regions associated with internalised cognition (Angular gyrus, ATL, Precuneus, and medial PFC) show tight regional co-ordination forming a distinct cluster from Amygdala and AITD.

Importantly, person-knowledge stores can be accessed in several other modalities (such as name & voice) in addition to face. Extended component of the person-perception system has been observed activated in the blind (Wang, Peelen, Han, Caramazza & Bi, 2016). However, whether the representation in the extended system changes according to the stimulus modality is unclear. It is important to establish whether knowledge is organised in an amodal modality independent way, or whether recalling knowledge is largely dependent on stimulus modality. Key feature of conceptual processing is invariance across stimulus modalities. To disentangle conceptual representation from perceptual processes, this chapter will address stimulus invariance of regional coordination of the extended system

In a PET study Gorno-Tempini et al. (1998) contrasted famous face and name processing. Expectedly, they found that some core regions, namely the FFA respond to faces but not names. When names were compared to meaningless strings of letters, they engaged portions of the posterior middle temporal gyrus and superior temporal sulcus, suggesting that these areas are name-specific. Analysis of famous faces + famous names – face and name control conditions revealed an extended network of regions that are involved in person knowledge regardless of modality of presentation. Authors found that regions implicated in the extended system of face processing – regions of the internalised cognition network, IFG and OFC respond to famous names as well. These results suggest that largely, regions involved in person-knowledge are at least cross-modal, i.e. respond to both face and name inputs. Within these regions, authors found that bilateral ATL and OFC are active during face and name processing but are more recruited during face-viewing than name reading. In a later fMRI study contrasting famous face and name processing, Nielson et al (2010) further highlighted similarities and differences between the two stimulus types. They again found modality-specific activations, regions in the temporal lobes selective for either name and face processing. Regions associated with internalised cognition (ATL, precuneus, AG, vmPFC) were not modulated by the stimulus type. IFG and OFC showed face preference effect - these are regions are more active during face perception, than during name reading, but are recruited for both. In addition, there were differences in auxiliary systems (regions not part of core or extended systems) recruited. Face processing was accompanied by pronounced activity in the medial part of the superior frontal gyrus.

Names recruited areas along the cingulate gyrus, post-central gyrus among others. How these distinct-yet-overlapping activations affect the organisation of knowledge within these regions is still unclear. Whether modality-specific activations accompanying face and name processing affect the relationships between extended system regions is unclear.

Here we investigate the role of stimulus in the organisation of person knowledge. More specifically, we investigate to what degree there is the *coordination* within extended person-knowledge system amodal. Stability across stimulus modalities would indicate a single distributed system accessed by both famous faces and names, providing insight into knowledge organisation. This comparison is needed because it is possible person-knowledge is encoded differently, depending on the stimulus modality. By closely following the methodology we developed in the previous chapters with a single pivotal manipulation - changing the stimulus from a person's face to a person's full name accomplish two main goals: 1) Replicate findings about the organisation of knowledge and regional co-ordination and 2) Generalize the findings and rule out perceptual effects.

Methods

Participants

31 participants took part in this study (mean age $M = 24.74$, $SD = 3.13$). All participants had normal or corrected to normal vision and were free from neurological disorders. Study was approved by the University of Trento Ethical committee. Participants were compensated for their time. Seven participants were removed from analysis: Two

participants did not complete the experiment due to technical issues with the stimulation computer, five participants were removed due to excessive motion in the scanner.

Stimuli

Stimuli referred to the same 40 celebrities were used as in chapter 2. Instead of presenting their faces, famous person's first and last names were presented on separate lines. Mean name length (first name + last name) was $M = 12.15$, $SD = 2.59$ letters. All experimental conditions were kept the same with one exception. "Full name" task in the previous experiment required participants to "recall persons full name". For the current experiment this was changed to "how common is this person's last name".

Data analysis

Preprocessing and GLM fitting pipelines were kept the same as in chapter 2. To compare organisation of knowledge and functional co-ordination only across the extended system ROIs, we restrict the ROI set by removing core regions - OFA, FFA, pSTS (bilaterally) from the analysis. We use similarity structures from the face experiment (regional and cognitive similarity matrices, averaged across participants ($N=20$) as a model (figure 3.3), and compare it against similarity patterns observed in the current experiment. Correlations between model and individual similarity patterns were then subjected to a one-sample t-test, to establish significance.

Results

Whole brain analysis

Contrasting all tasks over 1-back name matching control task revealed an expected network of activated regions. Within the extended system, most of the regions previously observed active during face-queued person knowledge are seen activated during name reading as well (see figure (3.1), also chapter 3, figure 2.6). Left ATL, precuneus, bilateral AG and prefrontal cortices are active during person knowledge regardless of input modality. Name reading specifically engaged areas left pre-motor cortex. We next extract task-specific response magnitudes from independently defined, face selective ROIs for further analyses.

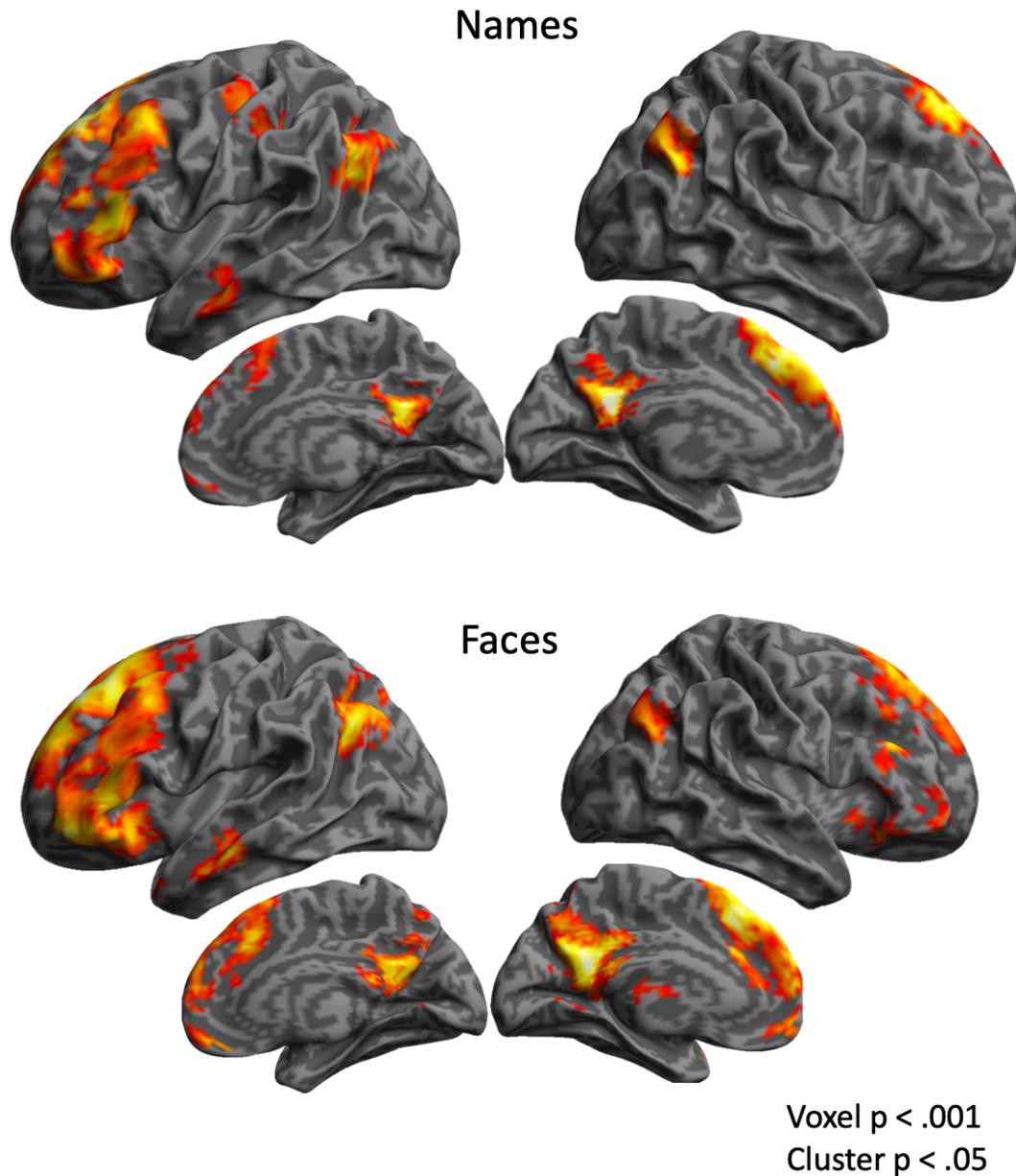


Figure (3.1). Regions involved in person knowledge when reading names (top) and viewing faces (bottom). Figure shows regions significantly active in all tasks > control contrasts. Control tasks were 1-back matching tasks with either name or face stimuli. Both experiments elicited broad clusters of activation, when access to person-knowledge is required.

ROI analyses

In chapter 3, to isolate cognitive response, we subtracted activity magnitudes elicited by perceptual 1-back matching task of famous faces, from the average activity produced when accessing person knowledge. This revealed robust activation of regions involved in person-knowledge such as regions of the internalised condition network, as well as left IFG and bilateral OFC. We perform the same analysis in the current experiment as well. Namely, we subtract response magnitude elicited by 1-back name matching task, from the average magnitude observed during person knowledge tasks. This allows us to compare regional involvement in person-knowledge, abstracted from perceptual stimulus effects, figure (3.2). Broadly, face perception recruits extended system regions to larger extent, as evident in the regional amplitude differences. Overall right hemisphere regions were engaged more than the left, $t(42) = 2.59, p = 0.013$. Individually, no single region showed face preference after correcting for multiple comparisons.

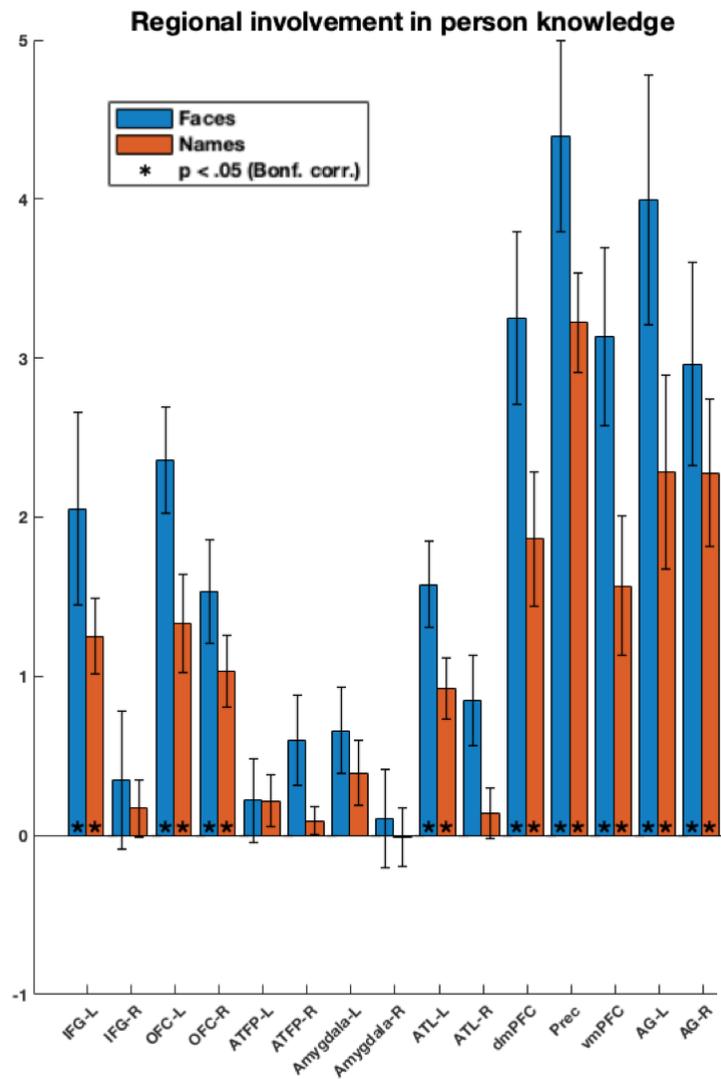


Figure 3.2. Extended system ROI involvement in person knowledge during face and name reading. Independently defined ROI active during face viewing tend to be active during name reading as well. Face viewing engaged regions more than name reading does.

Regional co-ordination stability across modalities

We first investigate regional co-ordination stability, when going across modalities. We use the average regional similarity matrix from the previous experiment as a model and correlated it against each subject's regional similarity matrix in the present experiment, figure 3.3. In this way we compare the regional coordination across tasks when we look at faces to the regional coordination profile when we read names. Results show strong similarity in functional connectivity patterns across experiments $t(23) = 10.30, p < 0.001$, indicating that overall co-ordination is largely preserved across input modality. We visualize the regional co-ordination structure, by clustering similarity patterns of the two experiments. Looking at clustered patterns of similarity, we see that regions associated with internalised cognition – precuneus, ATL, vmPFC, dmPFC and AG preserve shared similarity structure during name reading, consistent with findings from the previous experiment. Amygdala and AITF again demonstrate cognitive profile similarity and cluster apart from the internalised cognition regions. Comparing competing models of regional coordination we observe that AITF shares more response pattern similarity with the amygdala than the internalised cognition network $t(23) = 3.77, p < 0.001$.

Focusing on differences, we observe that IFG clusters with the internalised cognition network, while OFC clusters with amygdala and AITF. To see whether this effect is common to all subjects, we directly compared observed organisation against the model derived from the previous chapter. Namely, in this post-hoc analysis, we constructed an RSA model in which IFG couples with internalised cognition regions, and OFC couples with amygdala and AITF. We compare this model to the average regional coordination

matrix from the previous experiment. We observe robust model fit suggesting that IFG coordinates with the internalised cognition regions, and OFC coordinates with amygdala and AITFP $t(23) = 2.62, p = 0.015$.

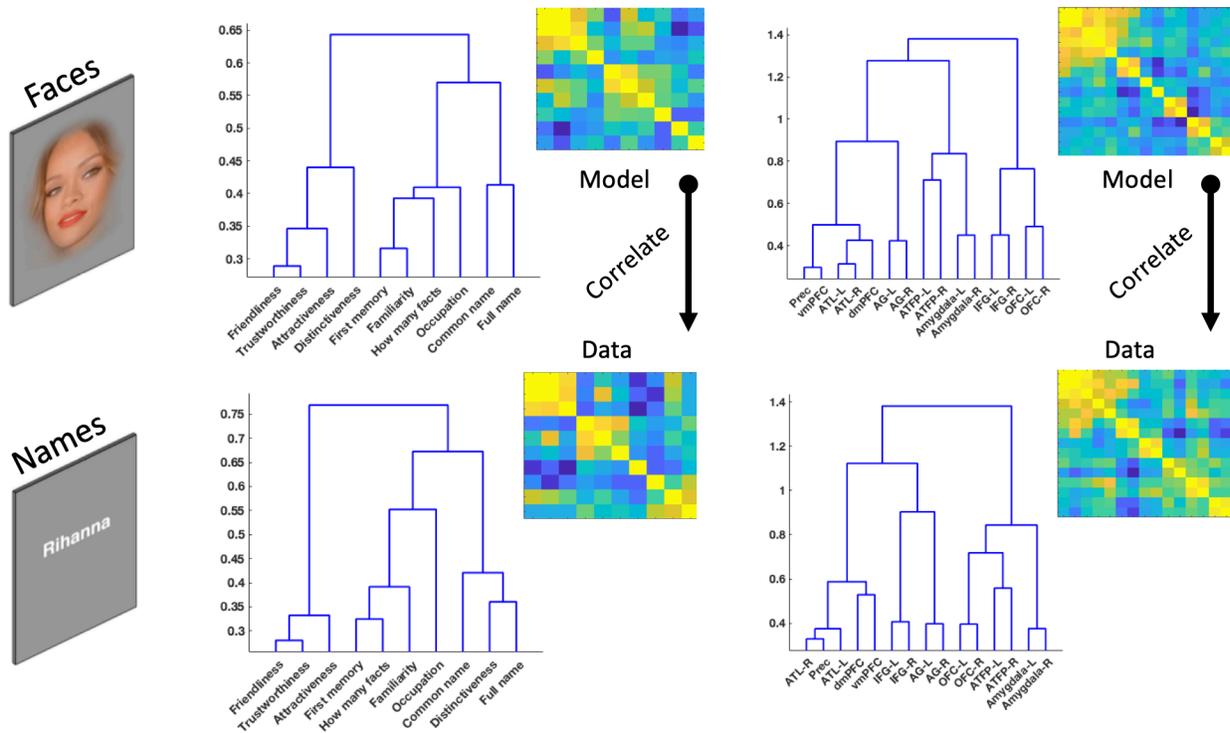


Figure 3.3. Cognitive and Regional organisation within the extended system regions involved in person knowledge. Top row: average clustering from the face experiment was used as a model and was correlated to the observed clustering during name reading. Bottom row: Results of cognitive and regional clustering during name reading.

Conceptual stability across modalities

Overall, cognitive structure remains largely the same, using the data from the previous experiment as a model we find significant correlation between task similarity matrices

within the extended system $t(23) = 5.33, p < 0.001$. This is contrast to the core system, where there was no similarity between face data model and name data for task similarity $t(23) = -0.69, p = 0.496$. Visualizing the cognitive similarity across the extended system ROIs in the name experiment, we see an overall similar structure like we did in the face experiment, with the exception that distinctiveness task clusters with nominal knowledge during name presentation. To confirm our previous findings that biographical knowledge is more similar to episodic knowledge than knowledge about names, we fit competing RSA models. Results replicate previous findings, $t(23) = 2.96, p = 0.007$.

Discussion

In this experiment we investigated person knowledge representation during name reading. We show that the same regions involved in person knowledge during face perception are also active during name reading. This attests to the cross-modality of the extended system and supports its role in knowledge, rather than perception. Regions showed a more pronounced effect during face processing, possibly because we used face sensitive regions-of-interest defined via faces > animals + tools localizer. Despite this, we demonstrate that person-knowledge encoding, and regional co-ordination remain stable regardless of stimulus type.

Differences in regional coordination

We describe effects of stimulus modality on regional co-ordination. Cognitive similarity between different regions involved in person knowledge remains largely stable. During both face viewing and name reading, regions involved in internalised cognition remain

tightly connected, suggesting that stimulus modality does not affect how these regions co-ordinate. Our finding that ATRP shares response similarity with the amygdala, together forming a distinct part of the extended person knowledge system, is replicated. In the absence of an active core system, IFG and OFC regions couple with different parts of the network.

IFG, previously implicated in guiding attention is seen to couple with internalised cognition regions. This pattern of results has been observed before. Humphreys et al (2015) observed IFG co-activate with the intrinsic semantic network specifically during semantic tasks involving reading. The activity in IFG attenuated when the stimulus was changed to pictures. Some accounts implicate IFG in guiding attention. In tasks where attention demands are focused inwards (such as creativity tasks), IFG has been observed to couple with the brain's intrinsic semantic network (Beaty et al. 2014). Our results are in line with these findings. Flexible connectivity profile of the IFG, coupling with perceptual regions during face viewing, and intrinsic semantic regions during reading is further evidence for its modulatory role in person knowledge.

Lateral orbito-frontal cortex (OFC) has traditionally been implicated in reward and facial beauty. During face viewing we observed that OFC couples with core perceptual regions, in line with accounts which found that attractiveness modulates perceptual, and lateral frontal regions (Chatterjee et al, 2009). It is consistent then, that in the absence of an active core system, OFC would cluster close to amygdala, a region known to be involved in facial attractiveness judgements (Winston et al., 2007). Throughout the years, however, role of OFC has expanded, Lindner et al. (2008) found increased OFC

activity during person comparison for both physical (height) and mental (intelligence) judgements. Such findings are inconsistent with purely reward/attractiveness account of the OFC, as it suggests a more complex role (mental processes unrelated to emotional valence).

Taken together, these results show that cognitive profiles of the extended system regions are stable across modalities. Regions involved in internalised cognition remain tightly coupled, amygdala and AITC form a functionally distinct part of the extended system. During face-viewing, we observe robust response-pattern similarity between IFG, OFC and the core system. Shifting the stimulus to names revealed previously overshadowed, more subtle similarity patterns between these regions and the rest of the extended system.

Cognitive Domain Similarity Structure is preserved across modalities

In chapter 2 we observed that different cognitive efforts (episodic retrieval, remembering biographical facts, social judgements, accessing nominal knowledge and physical judgements) rely on overlapping neural substrates. We showed that in the context of person-knowledge episodic and biographical domains are encoded similarly across the person perception network. Here we demonstrate that this cognitive structure is largely unaffected by stimulus modality. Episodic memories and biographical facts cluster together, during both face and name presentation replicating our previous finding that these cognitions rely on similar extended system regions. This suggests that the

extended system regions jointly encode stable representations of correlated cognitions, abstracted from perceptual features.

Judging facial distinctiveness, de-clusters from previously observed social-physical judgement cluster, suggesting that, at least some kinds of physical judgments are affected by stimulus modality. It's unsurprising that out of the cognitions tested, judging facial distinctiveness would be the task most sensitive to stimulus modality. Attractiveness (other exemplar of physical knowledge domain of tasks) is intrinsically linked to both physical features and social norms, possibly explaining its robust encoding across modalities. Prosopagnosic patients have troubles judging face distinctiveness but not attractiveness (Carbon, Grüter, Grüter, Weber & Lueschow, 2010), suggesting selective benefit that perceiving the face has on distinctiveness, but not attractiveness tasks.

Notable differences aside, both regional co-ordination and conceptual structure remain largely intact across the two stimulus modalities. Regional and cognitive similarity models built using face modality data show robust fit, when correlated with data from written word modality.

Conclusions

In chapters 3 & 4, we investigated person-knowledge organisation in core and extended person knowledge systems. In two separate experiments we looked inside the processing happening within the core and extended components of this system. By manipulating the aspects of person knowledge and the stimulus modality we were able

to parcellate different functional units of this coordinating system. Historically investigated in the context of face perception, the extended system subserving person knowledge shows robust cross-modal response. We demonstrate that the regions involved in internalised cognition form a distinct part of the extended system. These regions coordinate closely across different cognitions and stimulus modalities. Amygdala and AITD consistently form a distinct cluster of the extended system. Regional coordination of IFG and OFC are more dependent on the input modality. IFG couples with perceptual regions during face viewing and the internalised cognition network during name-reading, possibly reflecting attentional demands.

Overall, person-knowledge is linked and relies on shared neural substrates. Instead of being localized in individual regions, diverse kinds of person-knowledge seem to be distributed across ensembles of regions. Seemingly different forms of person-knowledge, such as biographical facts and episodic memories are encoded in overlapping brain regions. Diverse kinds of person-knowledge are accomplished by the coordinated activity of the extended system. Person knowledge representation is abstracted from perceptual features and is stable across stimulus modalities. Next chapter will investigate whether this holds true for our knowledge of object categories.

Chapter 5: Conceptual knowledge of object categories

Introduction

Previous chapters focused on specific type of knowledge - person knowledge. More specifically we investigated how different domains of person knowledge are organised across the core and extended system for person perception. We observed that different domains of person knowledge engage overlapping neural substrates. Here we will investigate whether these findings generalise to our knowledge of broader semantic categories and conceptual distance.

Object categories are encoded differentially across the cortex. Evidence from neuropsychological literature has demonstrated that focal damage to the brain can produce impairments accessing knowledge about specific categories of objects, often regardless of input modality (visual or verbal) (for a review see: Thompson-Schill, Kan & Oliver, 2006). Patients have been observed to have impairments to broad categories, such as only living things, or only non-living things (Martin & Caramazza, 2003). Other patients have even more specific impairments that only affect animal, fruit & vegetable or tool knowledge. (Thompson-Schill, et al., 2006). Broadly patients with prominent lesions in the temporal lobe, are more likely to exhibit impairments to living things, while non-living category knowledge impairments are associated with extensive fronto-parietal damage (Thompson-Schill, et al., 2006). Theories aimed at explaining where and how different concepts are encoded have not reached a consensus. It is thought that the stark distinction between the way object categories are encoded in the brain, is, at least to a degree, due to the different loading on attributes (also called features) they possess. Our experience with living things relies heavily on our sensory interaction with them,

making sensory features, such as their size, colour, smell, and taste - more important at differentiating them. In contrast, non-living things rely heavily on functional features - such as shape (“has a handle”) and intended purpose (“is it used for cutting”) (Warrington and McCarthy, 1983).

Despite observable impairments to category processing, observing category selectivity with fMRI/PET has been challenging (for a meta-analysis of natural objects and artefacts see Gerlach, 2007). A large-scale meta-analysis by Binder et al (2009) identified seven key regions reliably recruited during semantic processing, spanning parietal lobes (Inferior Parietal Lobule (IPL), angular gyrus), temporal lobes (posterior middle temporal gyrus (pMTG), parahippocampal gyrus (PHG)), frontal lobes (dorsomedial and ventromedial prefrontal cortices (dmPFC, vmPFC), inferior frontal gyrus (IFG)) as well as medial cortex (Precuneus (PC)). Despite wide-spread activation in response to semantic processing demands, little category specificity was observed. While portions of pMTG were more involved when accessing tool knowledge other concept categories (such as animals) did not rely on specialised neural substrates.

Commonly, individual studies investigating semantic processing observe activity in seemingly category-general semantic system. Devlin et al. (2002) reports three experiments focusing on five categories of objects: animals, fruit, tools and vehicles. Across the three experiments, authors reliably found a network of regions responding to semantic categories over a non-semantic baseline. However, no brain region was selective for any particular object domain (living, non-living).

Increasingly however, studies employing multivariate analyses observe that while category selective response may not be evident in the overall regional activity, the multivariate patterns of voxel activity are sometimes sensitive to semantic content (Bruffearts et al., 2013; Simanova, Hagoort, Oostenveld & Van Gerven, 2014; Fairhall & Caramazza, 2013). These studies report differences in fine-grained voxel patterns across semantic categories, such as mammals, birds & tools or in individual exemplars. These differences were evident in category general regions such as AG, IFG and pMTG (Simanova et al, 2014). Taken together, these results raise a question, how a largely category-general system, processing overlapping stimulus features, encodes category relationships. Whether knowing that spatula and a cat belong to different conceptual categories is encoded in a particular region or emerges as a product of the coordinated activity across multiple the semantic regions is still unclear.

Competing theories addressing knowledge organisation in the brain have been proposed. Some suggesting that regions process specific stimulus features and contribute to a single, high-level hub (Patterson, Nestor & Rogers, 2007). Other accounts do not assume a hub. Massively distributed account (Tyler, Moss, Durrant-Peatfield & Levy, 2000) suggests that object categories emerge from subtle differences in content and structure within regions, with no explicit category information in any region. In this account, the driving force to behind category grouping is their overlapping sets of semantic features.

In addition to investigating category sensitivity, Fairhall and Caramazza (2013) investigated regional sensitivity to semantic distance. They observed that six regions:

VTC, AG, latPFC, dmPFC, PC and pMTG sensitive to semantic categories. However, only PC and pMTG were sensitive to semantic distance between categories. These findings suggest that that some regions involved in concept processing might not necessarily be encoding conceptual distance.

Current approach

The current study will investigate the distributed nature of object category knowledge representation. Building on Fairhall and Caramazza (2013), we will investigate the relationship between neural representation and semantic distance. Crucially, we will extend this approach to test the sensitivity of the overall semantic network to semantic distances. Specifically, we will ask whether there is more information about semantic distance arising from the coordinated network activity than is contained in individual regions alone. We will start by building a word co-occurrence based semantic reference model by clustering a large corpus of text to identify naturally occurring concept categories. Distributed accounts (Tyler, Moss, Durrant-Peatfield & Levy, 2000) predict that information about semantic distances would be visible at the network-level, whereas semantic hub theory (Patterson, Nestor & Rogers, 2007) predict this knowledge to be localized in a few key regions.

Method

Participants

Twenty right-handed, native Italian participants (mean age $M=25.8$ $SD=4.47$, 11 females) took part in the study. Participants had normal or corrected to normal vision

and were free from neurological disorders. All participants gave informed consent and the study was approved by the university of Trento ethical committee. Two subjects were removed from the analysis due technical problems during fMRI acquisition.

Stimuli & model

To sample naturally occurring semantic categories, we looked at how words cluster in written text. In a large corpus (wikipedia) we counted the co-occurrence of frequent, concrete nouns (1481 nouns, referred to as 'targets') with common verbs, adjectives and nouns (i.e nouns not already present as 'targets') - referred to as features (6773 features in total). Word occurrences were log transformed, and the correlation distance $(1-r)$ between targets was taken as a measure of word relatedness. We visualised word similarity by computing a dendrogram (1481 items, correlation across 6773 features), see figure (4.1). Word embedding similarity analysis captures well known categorical divisions, such as living-non-living distinction (Warrington & Shallice, 1984). Zooming in, within the living thing cluster we see that food related items are represented differently from other living things. Within non-living things clothing items and musical instruments are represented differently from the rest, in line with previous work using semantic feature production (Cree & McRae, 2003). Hypothesis raised by McRae and Cree (2002) that vegetables and plants would cluster differently because of their marked difference in ecological salience can be seen confirmed in the clustering. We used this word clustering to select categories to be used in the experiment. Some categories were too mixed to ascribe a label that fits all items, some categories were unintuitive when paired

with a typicality task. Final stimulus set contains 17 categories (one category was removed from the analysis, see results: Removed category), 24 items each. English words were translated to Italian. The control condition (1-back matching task) was performed on Italian function words (matched on word length to experimental stimuli).

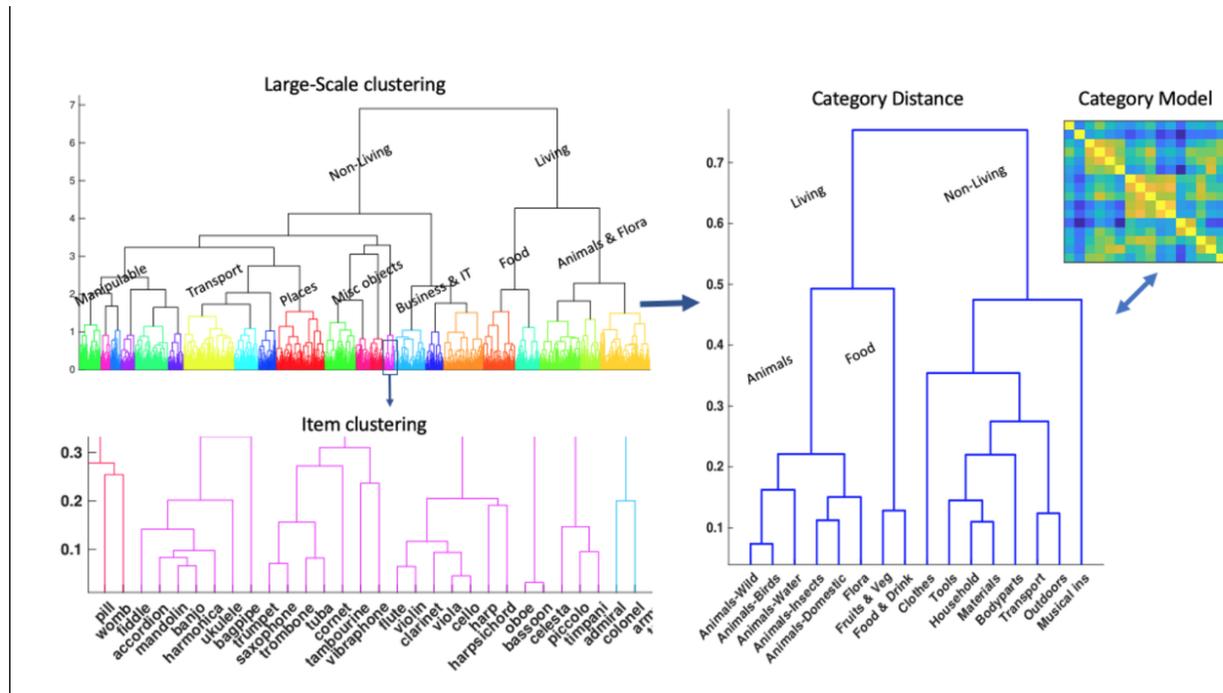


Figure 4.1. Corpus clustering procedure used to derive a semantic mode. A) Clustering of all 1481 items shows naturally emerging conceptual structure. Known division such as living and non-living can be readily seen. B) Zooming in on a representative cluster, we see that conceptually similar items, such as musical instruments, cluster together. C) Averaging across category exemplars we demonstrate category relatedness (dendrogram) and the model (matrix) used.

To get a measure of how categories are related we average the feature vectors belonging to the same category (e.g. for domestic animals that would be averaging

'cat','dog', etc.). Word embeddings for 24 items for each category (24x6773) we averaged into a 1x6773 co-occurrence vector, describing the category. We then correlated category vectors to build a 16x16 category similarity matrix. It is important to note, that the importance of semantic model lies not with categorical differences (animals vs food), but with the quantified degree of relatedness (how different are animals from food).

Procedure

Stimuli were presented in a blocked design. Each block started with 4s of fixation, followed by the name of the category (e.g. "Household items") for 4s, followed by an additional 2s of fixation cross. Then a stimulus from the category (e.g. "Spatula") was presented in the screen for 2s, during which participants were asked to read the word and rate how typical it was or the given category (How typical is spatula of a household item). Participants rated the items on a scale from 1 (low typicality) to 4 (high typicality) using the button boxes provided. Blocks consisted of eight stimuli presentations, each category block was presented three times during the experiment so that 24 items were presented once, and all stimuli that participants saw were unique. Mean word length was $M = 5.5$, $SD = 1.78$ letters, ranged between 3 and 12 letters, and did not differ significantly between any pair of categories (all $p > .001$, independent samples t-test).

Data acquisition

MRI data was acquired using SIEMENS Prisma scanner (field strength - 3 Tesla) with a 64-channel head coil. Total of 720, T2* weighted EPI scans (voxel size - 3mm, isotropic) were acquired during the experiment. The experiment was split into 2 runs, each lasting 12minutes (360 volumes). Functional scans were AC/PC aligned, and 30 axial slices (gap - 0.6mm, TR = 2000ms, TE = 28ms, image matrix = 64 64, flip angle = 75, 108mm brain coverage) were collected for each volume. To aid coregistration, for each subject we also collected a high resolution T1 weighted MPRAGE (1mm, isotropic) anatomical scan (208 sagittal slices, image matrix = 288mm 288mm, flip angle 12, TR - 2140ms, TE - 2.9ms). Stimuli were presented on a 42", MR-compatible Nordic NeuroLab LCD monitor positioned at the back of the magnet bore that participants saw through a mirror in front of them. Stimuli were presented using a custom PsychToolBox 3 script running on top of Matlab R2017b.

Data analysis

Data were pre-processed with SPM12. Functional images were realigned, grey matter segmented, warped into common space and smoothed with 8mm FWHM kernel. Subject specific response estimates (beta weights) were derived by fitting a general linear model (GLM) to the data. GLM was estimated using a blocked design of 24 unique stimuli per each category and each block consisted of 8 stimulus repetitions. Each block was presented 3 times during the experiment. Estimated motion parameters from the realignment procedure were included in the model as regressors of no interest. To

isolate regions sensitive to semantic content we compute a contrast the average of all categories > function word 1-back matching control condition. We hypothesise that regions sensitive to semantic content might not overlap with regions sensitive to semantic structure. An omnibus, one-way ANOVA test was also computed, with 16 categories serving as factors. ANOVA test highlights regions that demonstrate the most change across categories. Whole brain analyses served as bases for ROI selection.

ROI definition

Region-of-interesting definition procedure was the same for both contrasts (all vs control and ANOVA), figure 4.3. We selected regions of potential relevance based on previous literature (see Binder et al, 2009). Regions that were active but were not associated with semantic processing (e.g. activations in the striatum) were ignored. After identifying region of interest, we located nearest local maximum and drew a 6mm sphere around it. For each subject, category specific beta values were extracted for each ROI, for each run, averaged across runs and subjected to further analyses (see: multivariate analyses). Complete list of ROIs used can be seen in appendix 1.

Multivariate analyses

Multivariate analyses leveraged the patterns of activation either across voxels within a ROI (Representation Similarity Analysis (RSA)), or across a network of regions (netRSA), figure 4.2. For both RSA and netRSA procedures we measured pattern

similarity for each pair of categories. Similarity measure in a single region (RSA procedure) was a Pearson correlation between voxel patterns for all categories (pairwise). For netRSA we averaged across ROI voxels and for a given category pair, computed similarity of categories across regional activations (network pattern correlation). Both RSA and netRSA procedures result in a 16x16 category similarity matrix. The similarity matrix for each subject is then correlated to the semantic model (see: stimuli & model) and the correlations are subjected to a one sample t-test (null hypothesis $r=0$) to establish statistical significance (correlation between model and regional or network representational spaces are referred to as RSA fit and netRSA fit, respectively). To test whether information is predominantly distributed or localized, we use a stricter test and directly compare netRSA fit version average model fit in individual regions (null hypothesis that netRSA = average regional model fit).

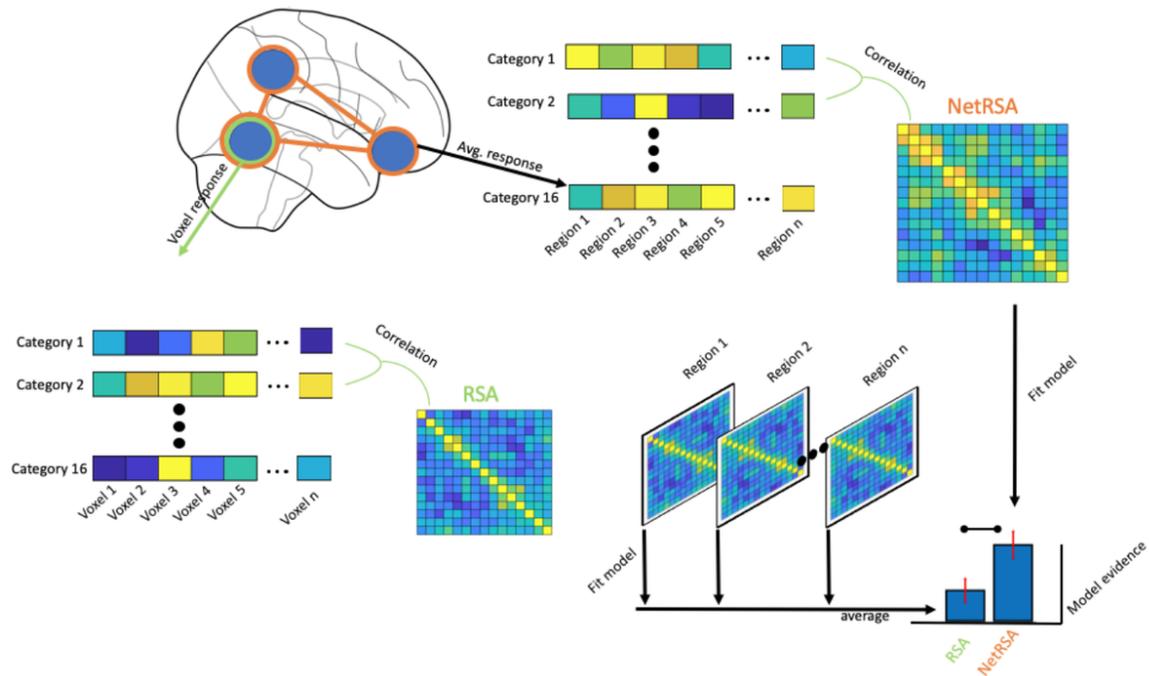


Figure 4.2. Multivariate analyses procedures. Regional model fit (RSA) was averaged across regions and compared to network-level model fit (netRSA).

To directly compare the two methods, we average across RSA fit from different regions, and compare against netRSA fit, in a paired samples t-test (see results: Information content at the Network-level vs Region-level). By Comparing average RSA fit against netRSA fit, we isolate the information that emerges from inter-regional coordination, after accounting for isolated regional information content. Lastly, we investigated fine-grained representation (both within individual regions and across the network), by reducing both the model and the data (n=16) to only include living (n=8) and non-living (n=8) categories, in separate tests. Investigation of fine-grained representation accounts for large-scale differences (broad differences between living and non-living categories) potentially influencing the model fit.

Results

Removed category

Whole brain analyses revealed that one category out of 17, consisting of leisure related concepts (“book”, “ball”, “song” etc.) produced unexpected brain activation patterns, affecting the whole brain analysis results. ANOVA contrast showed unexpected range of F values (up to 60) across 17 categories. After removing leisure category, F values fit within an acceptable (from 0 to 7) range. We removed leisure category from all subsequent analyses.

Behavioral Results

Across 16 experimental categories, there were significant differences in reaction times, $F(15,225) = 3.952$, $p < .001$. On average, however, there were no difference between living ($n=8$) / non-living ($n=8$) category domains, $t(17) = -0.62$, $p = 0.5453$. Typicality ratings differed across 16 categories, $F(15,225) = 15.659$, $p < .001$. Items in domestic animals category were rated as being less typical than other categories. Differences in behavioural responses were anticipated during experiment design and steps taken to elucidate whether they affected variables of interest are discussed further (see below: response similarity does not explain brain data).

ROI selection

Highlighting regions sensitive to semantic content, we observe broad clusters of activation spanning temporal, parietal, frontal & medial lobes, consistent with findings from previous studies investigating semantic processing (all categories > control condition, $p < .001$), (figure 4.3, red). Next, we isolate regions sensitive to change across categories, as opposed to change relative to control condition (Figure 4.3, green). We computed an ANOVA contrast across 16 categories. Whole brain maps served as basis for ROI selection. Two different sets of ROIs were analysed, corresponding to two different whole brain contrasts (all categories > control & an omnibus ANOVA). Categories > control contrast were designed select regions common to all categories, while ANOVA ROIs highlighted regions specialized to only some categories.

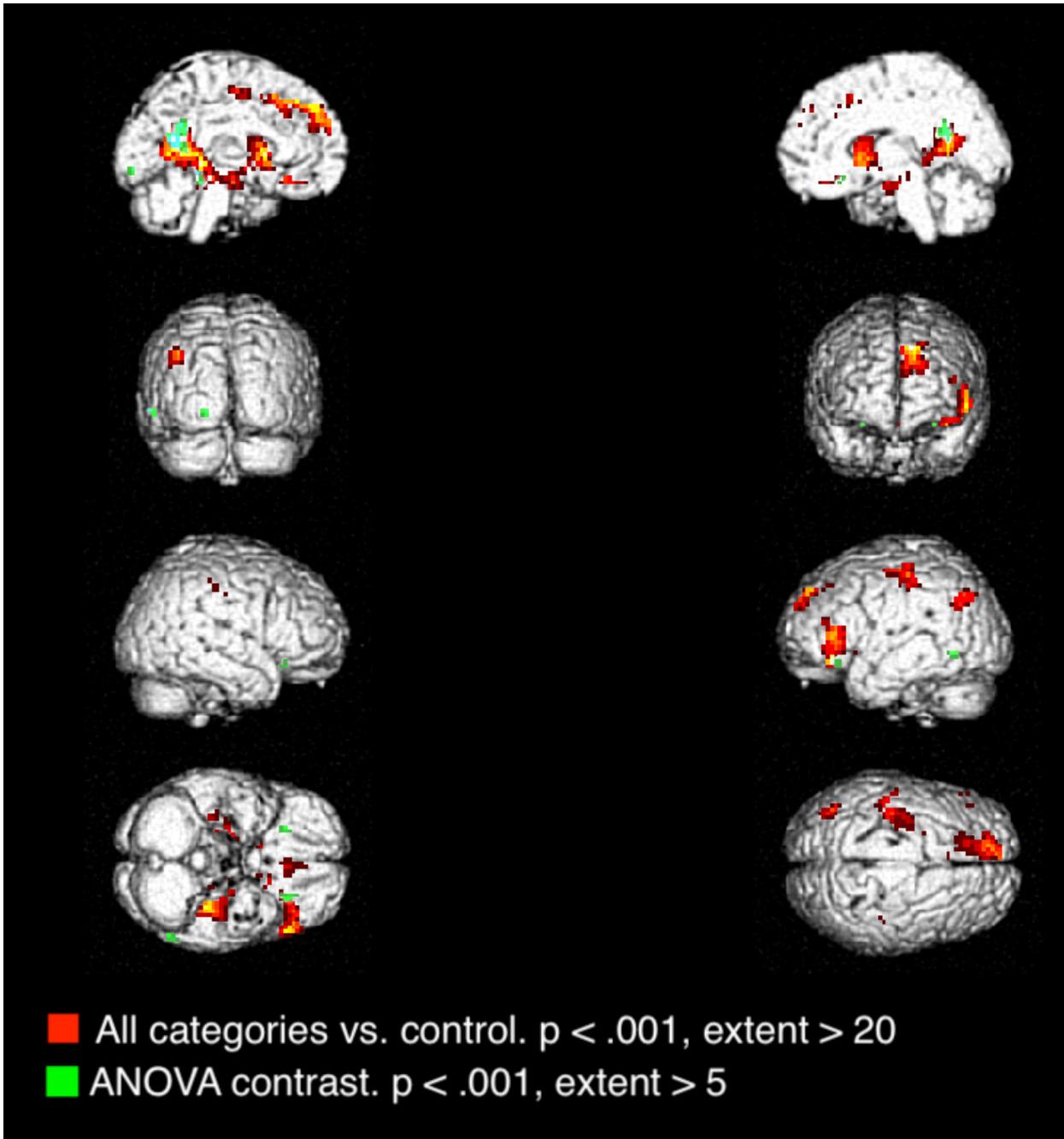


Figure 4.3. Whole brain analyses. All categories vs control contrast (red) shows expected, broadly distributed set of semantic regions reliably activated by all tasks. In contrast, ANOVA (green) shows regions demonstrating most variation across tasks. Peaks from both contrasts served as ROI centres for multivariate (RSA and netRSA) analyses

Regions common to all categories do not encode category relationships

We first investigate sensitivity to semantic distance within regions common to all categories (all categories > control ROIs). There was no concordance with the model either at the network level, netRSA fit - $t(17) = 1.55$, $p = .139$, or at a single region level, (RSA fit, all $t < 2.482$, all $p > 0.024$ (unc.)). Further analyses focused on ROIs that demonstrated sensitivity to change across categories (defined via omnibus ANOVA).

Sensitivity to semantic distance in individual regions

Having seen that regions commonly activated for all categories (all categories > control contrast) are not sensitive to semantic distance (individually, or as a network), we shift our focus to a different set of ROIs, defined via an omnibus ANOVA test. This highlights regions that exhibited most change across semantic categories (figure 4.3, green). Here we report an investigation of how these regions behave at the individual level. Fitting model individually to semantic regions (RSA fit), we see that some regions contain rich information about semantic relationships. Three regions survived multiple comparisons, left OFC $t(17) = 5.01$, $p < 0.001$, left PHG $t(17) = 3.81$, $p = 0.001$ and left pMTG $t(17) = 3.34$, $p = 0.004$. This suggests that there is substantial amount of semantic information contained locally (in individual regions).

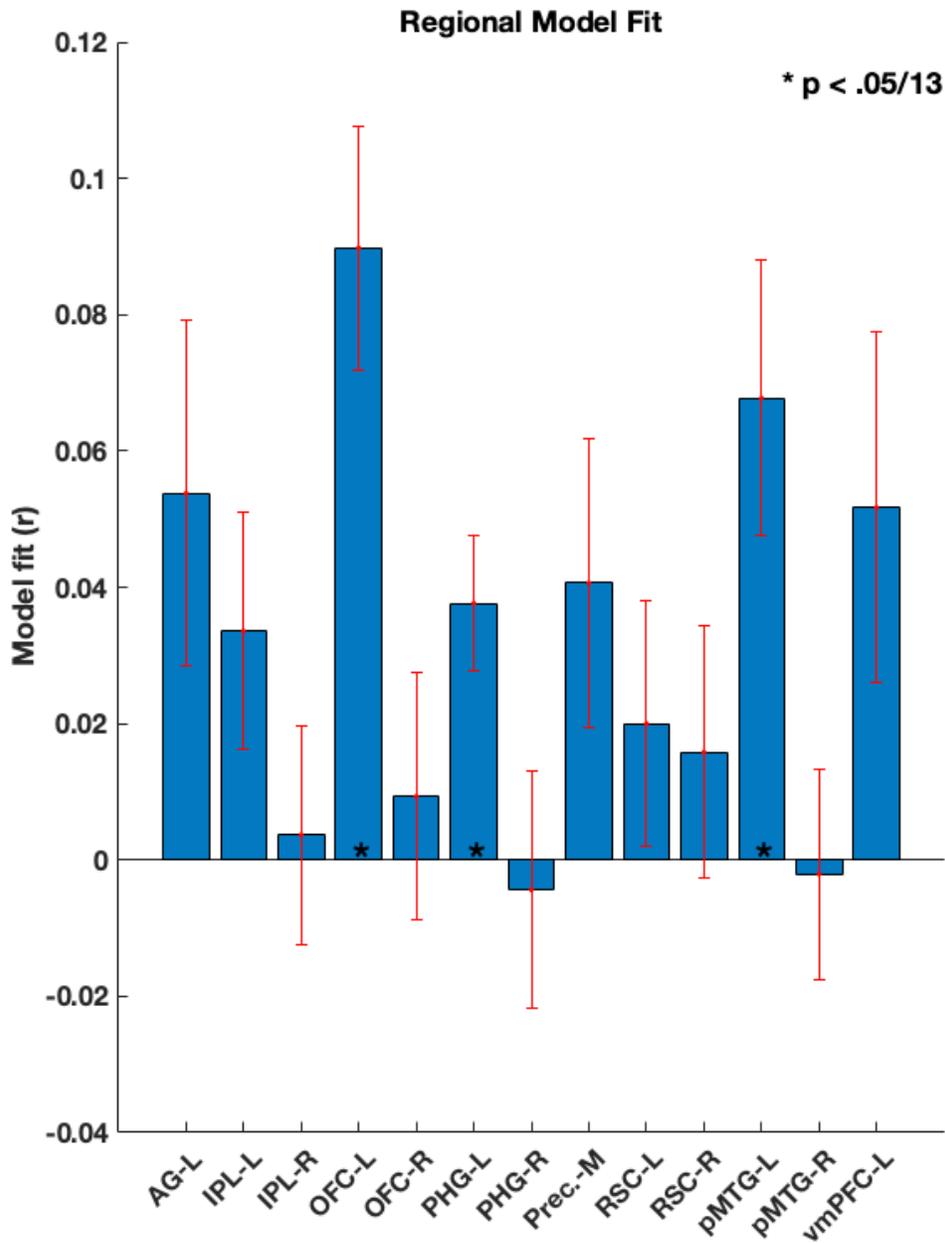


Figure 4.4. Regional model fit in regions significantly active in ANOVA contrast. Voxel activity patterns in these regions were compared to semantic model for each region individually. Activity in OFC, PHG, and pMTG significantly correlated with the semantic model (star denotes $p < .05$, corrected for multiple comparisons).

Information content at the Network-level vs Region-level

After investigating local information content, we focus on inter-regional behavior, by looking at these regions as a network (netRSA procedure, figure 4.2). netRSA revealed that network patterns reflect canonical similarity relationships $t(17) = 3.82, p = 0.0014$. Further analysis, directly contrasting information contained in individual regions (average of individual region model fit, figure 4.4) versus information contained across the network (netRSA fit), revealed a significant information gain when considering network as a whole, instead of looking at the sum of its parts $t(17) = 2.40, p = 0.0283$. Introducing across-region effects, by averaging regional RDMs results in information content equal to that observed in netRSA, difference between averaged RDMs > netRSA $t(17) = 0.41, p = 0.6876$. Taken together these results suggest that sensitivity to semantic relationships between categories is reflected in the network activity i.e. as the interaction between brain regions, more so than local activity.

Response similarity does not explain brain data

It is possible that systematic differences in response patterns across categories influenced the pattern of cortical response. Here we investigate whether our results were potentially confounded by variations in typicality ratings or reaction times (RT). For each subject we constructed separate RDMs based on RT and typicality ratings. We then compared them to netRSA and individual region similarity matrices to see whether reaction times or rating had an effect on the brain's representational structure. Behavioural matrices did not significantly correlate with network-level RDM ($t(17) = -$

1.78, $p = 0.0930$) or individual region RDMs (typicality fit: $t < 1.70$, $p > .107$, RT fit: $t < 1.84$, $p > .083$).

Regional vs Voxel coordination

Is the information about category distinction encoded between regional or voxel interactions? To answer this question, we investigate whether RSA can be further improved by collapsing across ROI voxels to form a combined ROI (583 voxels). We then performed the same RSA procedure as we did for individual ROIs (figure 4.2). Comparing voxel concatenated netRSA vs regionally averaged netRSA revealed no difference in model fit $t(17) = 0.54$, $p = 0.595$, and high correlation between RDMs ($M=0.908$, $SD = .057$) suggesting that regionally averaged network activity captures as much information as fine grained voxel patterns.

Fine-Grained representation of categories

Lastly, we sought to investigate whether network-level activity is enough to allow discrimination of fine-grained category distinctions. We separately computed semantic template fit for living and non-living kinds. There was no information that would differentiate between items within living kinds $t(17) = 1.26$, $p = 0.2256$ or within non-living kinds $t(17) = 1.08$, $p = 0.2941$ encoded at the network level. Individually, none of the three regions were able to differentiate between different living kinds (all $t < .79$, all $p > .219$). Left PHG gyrus was sensitive to differences between non-living things $t(17) = 2.22$, $p = 0.02$, but pMTG and OFC were not ($t < 1.44$, $p > .08$).

Discussion

In the current experiment, we investigated the distributed nature of object category knowledge. Extending the approach that we applied to person knowledge, we sought to investigate whether knowledge about conceptual distance of object categories is distributed or localised. Contrasting local versus network-level activity we observed that there is more information about conceptual distances at the network level, than is stored locally.

More specifically, by looking at brain pattern similarities, we were able to capture the cognitive processing happening within individual regions, as well as network overall. By then comparing representational spaces with a corpus derived semantic model we measured the sensitivity to semantic distance. We demonstrated that regions contribute to a robust network level representation. This work highlights the importance of focusing on cross-regional interactions, in addition to regional function.

Model

Our investigation is only as good as our model. To look for brain regions reflecting semantic distance we built a reference model to compare against. We sampled a broad set of concepts and chose word co-occurrence as a measure of overall similarity because it allows us to identify groups of nouns belonging to different contexts. By then clustering bottom-up, i.e. allowing item similarity form homogenous groups - we effectively sample naturally occurring conceptual categories. This approach proved

robust - the model contains well-known divisions - the main one separating living/non-living kinds. Within the living domain of categories, we observed a known division between animals and vegetables. Interestingly, composite foods (such as cake, and tea) share context similarity with living things, this, while factually untrue - is a feature observed in both feature generation tasks, as well as neuropsychological literature. Patients with impaired living kind knowledge usually have impaired non-living food knowledge (Cree & McRae, 2003). The relative degree of semantic relatedness between categories in our model is also consistent with previous literature on semantic feature production (vegetables cluster late with food, musical instruments cluster late with non-living things, body parts cluster with non-living (Cree & McRae, 2003)) suggesting that we have faithfully captured similarity relationships between our concept categories.

Regions not sensitive to semantic distance

Focusing on regions more active during semantic access than the control condition, we observe expected network of semantic regions. All categories, overall, recruited a well-known semantic network of regions, spanning temporal, parietal, frontal and medial lobes of the brain. This validated the efficacy of our task in eliciting semantic activation and served as a starting point for a targeted analysis of regions involved in semantic cognition. Investigation of this network revealed no sensitivity to semantic distance. Neither activation patterns within individual regions, nor combined activity across a network reflected semantic distance between categories. One explanation of this pattern of results can be that focusing on peaks common to all categories imposes a constrain

that dilutes specificity and rewards domain-general voxels. This is consistent with some accounts of neuroimaging of semantic memory that observed no categorical differences (such as Devlin et al., 2002), however others have (e.g. Fairhall & Caramazza, 2013b). Given this discrepancy in the literature, we shifted our focus and used an omnibus ANOVA test to highlight regions most sensitive to change across categories, instead of those responding to most categories. This targeted investigation revealed small foci of activation, often close to the previously described peaks. We interrogated the coordinated and isolated activity of these regions to see whether the change across categories that they exhibit corresponds to conceptual distance.

Statistical power

All categories > control task contrast revealed a network of regions sensitive to semantic content. These regions were significant at acceptable thresholds (peak $p < .001$, $k > 20$). We do not think that increasing statistical power would have drastically changed the location of these regions. Even given the acceptable locations of these regions, neither fine-grained voxels patterns, nor network level patterns reflected semantic distance, as measured RSA. On the other hand, defining regions via omnibus ANOVA test yielded much smaller activation foci. While it is possible that such effect could be due to statistical power, it can also signify reliably similar activation magnitude differences for different categories in our experiment. If regions chosen via ANOVA contrast were not involved in semantic processing one would expect them to not fit the semantic model tested. Instead, modest ROIs revealed by ANOVA showed were robustly sensitive to

semantic distance at both regional and network levels of analysis. Taken together, statistical power seems to have been adequate in this experiment.

Network level representation of semantic distance

By focusing on peaks sensitive to change across categories we highlight a network of regions that contribute to a robust, distributed representation of semantic content. We then directly compared representational content of individual regions, and the network as a whole. This revealed that more information about semantic distance emerges as a product of coordinated activity than is contained within individual regions. These results support distributed theories of conceptual organisation. We further demonstrate that network pattern devoid of local information (netRSA) captures as much information about conceptual distance as a measure that includes both local and distributed information (voxel concatenated RSA). This suggest that semantic distance of object categories is encoded in the macroscopic scale of regional interactions.

Different theories have been proposed to explain how the information is passed through a cortically distributed semantic network. Hub-and-spoke model (Rogers et al, 2004) proposes a single hub within which modality-dependent representations converge. Distributed accounts (Tyler, Moss, Durrant-Peatfield & Levy, 2000) propose that categories can emerge without an explicit hub. Anderson et al. (2018) demonstrate the correlated nature the semantic system. They observed that the representational spaces of semantic regions are highly correlated, and elaborate sentence meaning is

jointly encoded across groups of brain regions with no convergence hub. Our findings fit within a distributed representation framework. While we observe that some key regions encode category relationships, we demonstrate that information encoded at the network-level, (i.e. as a product of co-ordination) is significantly more robust.

Regional sensitivity to semantic distance

We uncovered three regions, robustly sensitive to semantic distance. Left lateralized, parahippocampal gyrus (PHG) posterior middle-temporal gyrus (pMTG) and orbito-frontal cortex (OFC), showed significant fit with the expected distance model. All three regions have been previously reported to be active during semantic tasks (Binder, 2009). Sensitivity to semantic distances is an expected property of a high-level convergence-hub region. We demonstrated that activation patterns in these regions reflect semantic distance. However, caution in interpretation is needed; Whether these activation patterns reflect conceptual distance or property (feature) processing is still unclear. Within domain differentiation is a strict test of knowledge representation. If a region houses robust semantic representations, it should reliably differentiate not only between conceptually different items (knowledge that “cat” and “spatula” belong to different domains of knowledge), but between relatively similar items as well (knowing that “cat” and “bird” are conceptually different animals). Here we discuss evidence of semantic processing observed within each region and weigh it against alternative explanations.

When we tested the differences between concepts within a macro domain (within living and within non-living separate) activity patterns in the left PHG were indicative of

conceptual non-living concept distance. Activity patterns within OFC and pMTG showed no sensitivity to within domain distance. PHG has been observed to be involved in entity-specific semantic access. More specifically, PHG robustly responds to scenes & large objects (Bi, Wang & Caramazza, 2016). Fairhall, Anzellotti, Ubaldi & Caramazza (2014) observed cross-modal activation in this region, favoring inanimate place concepts. OFC and pMTG regions also did not show expected sensitivity. Our results showing sensitivity to semantic distance in pMTG replicates previous findings. Fairhall & Caramazza (2013) first observed robust sensitivity to semantic distance in this region suggesting a central role in semantic processing. We did not observe within domain sensitivity in the pMTG. Whether the pMTG is sensitive to subtle within category differences remains to be established but accumulating evidence suggests its involvement in broad knowledge representation with a possible preference for tool knowledge (Binder, 2009). Activity in OFC during semantic tasks is often reported but often not discussed. Traditionally ascribed roles of emotion regulation and reward processing (Bechara, Damasio & Damasio, 2000) it's possible that activity in the OFC represents the reward value of the stimuli used. Cree and McRae (2003) showed that some categories comprising the living domain (animals, vegetables & foods) rank highly on the taste dimension, whereas non-living things rank low, offering a non-conceptual explanation of observed activity in the OFC. Taken together, when we cautiously interpret nature of category organisation in these regions, our evidence does not out-weight alternative explanations of processing within these regions.

In summary, this chapter investigated the distributed representation of object categories. To relate the findings to previous chapters - when accessing person-knowledge, category-general regions co-ordinate to jointly encode information. Person-related information is represented in shared neural substrates and differences between them are reflected in different network-level patterns of activity. In contrast when interrogating conceptual knowledge of object categories, we observed that regions common to all categories do not contribute to a significant network-level representation. Nor do they contain information about conceptual distance individually. When we investigated regions sensitive to change across categories (ANOVA contrast), we observed a network of regions contributing to a distributed representation of conceptual distance. While regions involved in person knowledge provide overlapping contributions, knowledge about the conceptual distance of object might be more sensitive to distinct contributions from specialized regions.

Chapter 6: General Discussion

Aims of the thesis

This thesis asked whether semantic knowledge is represented in the coordinated activity across networks of interacting regions. The question was motivated by knowledge that semantic processes are broadly distributed across the cortex. Diverse kinds of thinking engage both sensory/motor regions associated with perception and amodal regions, such as the ones involved in internalised cognition. Little is known about how the coordinated activity across regions gives rise to the wide repertoire of human conceptual knowledge. In a series of experiments, we investigate the representational structure not only of individual regions, but of networks as a whole. Aided by modern computational tools, we can get a glimpse at the coordinated processing across brain regions, as well as the structure of different cognitions. We demonstrate the correlated nature of knowledge representation - diverse cognitive efforts rely on overlapping neural substrates. We also highlight that regions do not perform functions in isolation – most-all cognitions recruit multiple brain regions. We conclude that representation of knowledge is broadly distributed across brain networks.

Person knowledge

Keeping track of people in our lives is a complex task. Each person that we know has a wealth of associated information such as their name, job, shared memories and social traits. How the brain organises and keeps track of this information is a fundamental question in social neuroscience. Early accounts postulated a special cognitive

mechanism for recalling someone's name, and that the rest of person-knowledge comes from a singular cognitive system (Bruce and Young, 1986). Later accounts, incorporating neuroimaging evidence, highlighted that different person knowledge aspects rely of different neural substrates. Haxby, Hoffman & Gobbini (2000) described a distributed neural network, comprised a core element involved in perception, and an extended system involved in person knowledge. The model also suggests that different cognitions rely differentially on extended system regions. For example, faces associated with strong emotions engaged the amygdala, while accessing episodic memories engaged regions involved in theory of mind (Gobbini & Haxby, 2007). These complex tasks are unlikely to accomplished in individual regions alone. Further research demonstrated that all regions of the extended system are recruited regardless-of and even in absence-of an explicit task (Todorov, Gobbini, Evans & Haxby, 2007). Together the results suggest a complex and correlated nature of person knowledge. In a series of experiments, we interrogated the structure of person knowledge and the brain mechanisms responsible.

Behavioural indices

Investigating behavioural indices of person-knowledge we start to observe that some cognitive efforts are similar to each other, and different from others. When we explicitly asked participants to arrange person-knowledge exemplars according to their similarity, we observe that social and perceptual tasks are rated being subjectively similar, and different from the ones that required remembering biographical facts, names, or memories. Subjective similarity between retrieval of episodic memories and biographical

facts put a well-known distinction between these two domains (Tulving, 1972) into a broader perspective. We observe that the largest distinction is between trait (social or physical) judgements and other forms of knowledge (episodic, semantic and nominal). We go on to show that some this structure can be recovered from implicit measures, patterns of response ratings provide insight into this latent structure of cognitive efforts. To investigate whether behaviorally similar tasks share similar neural mechanisms, we further investigated the neural activity associated with these cognitions.

Person-knowledge network organisation

When making sense of the world around us, the brain processes perceptual and semantic attributes separately. When perceiving faces, regions of the core perceptual system (OFA, FFA, pSTS) coordinate in analysing the visual features of the face culminating in face identification (Haxby, Hoffman & Gobbini, 2000). An ensemble of extended system regions works together to retrieve and integrate diverse information about the person.

By looking at patterns of similarity across brain areas we investigate how the coordinating brain regions in the core/extended person perception systems encode person knowledge. Building on previous findings we demonstrate that subjective grouping of tasks is largely replicated using multivariate brain analyses. We observe that, in the context of the broad repertoire of person-knowledge, tasks cluster into socio-perceptual judgements, episodic-factual retrieval and nominal knowledge. We highlight that task differences arise not from individual regions activity, but rather from subtle cognitive tunings and regional coordination. Focusing on cognitive tuning, we observed that most-all regions are involved

in most-all cognitions, with subtle variations between them. For example, while ATL is active for all cognitions, relating tasks to one another we observe that it is preferentially recruited by social domain of tasks. Similarly, Angular gyrus contributed to all cognitions, but is tuned towards episodic memories.

To ensure that our findings were robust, and to demonstrate input modality invariant knowledge representation, we additionally replicated the experiment using names, instead of faces, as a stimulus. Previous research has shown that the extended system regions, associated with retrieval of person knowledge, are engaged both by faces and by names (Gorno-Tempini et al., 1998; Nielson et al., 2010). Whether stimulus modality affects the representational content and functional coordination within these regions remained unclear. This second neuroimaging experiment allowed us to demonstrate the modulatory role of the IFG, which co-ordinates either with the core perceptual regions during face-viewing, or intrinsic cognition network during reading. This possibly reflects attention orientation. When viewing faces, more attention is directed towards the stimulus, while during name reading, attention is focused on internal stores. We also observed some notable stability when going across stimulus modalities. The internalised cognition network remains tightly correlated regardless of stimulus modality, suggesting that the representations within are amodal. These results contribute to an emerging consensus that these regions form the brain's *general semantic network* (Binder & Fernandino, 2015). Originally termed the default-mode-network (DMN) - defined by its task-deactivated characteristics (Raichle et al., 2001), recent evidence has implicated DMN in broader semantic processing, namely internally focused cognition.

These regions are frequently recruited together, suggesting coordinated processing. We expanded on these results, showing that during access to person-knowledge, these regions interact as a functional unit, characterised by shared response similarity across cognitions. Internalised cognition regions had a different response profile from other extended system regions, such as the amygdala and the AITC. During name reading, we observe increased co-ordination of these regions with the IFG, consistent with previous accounts (Humphreys et al., 2014). Coordination between internalised cognition network and IFG is both stimulus dependent (present during reading, but not picture viewing) and specific to semantic content (not present during meaningless word reading). To summarize the experiments on person knowledge, we demonstrate the complex and correlated way that person-knowledge is encoded. We show that relationships between different cognitive efforts as seen in subjective ratings and correlated patterns of responses can be traced back to shared neural mechanisms. Person-knowledge is stored in overlapping substrates within the extended system for person perception. These regions individually are involved in all person-knowledge efforts. Regions within the extended system coordinate to jointly encode diverse kinds of person-knowledge, abstracted from perceptual features.

General Semantic Object-Knowledge

Accessing semantic representations reliably engages a broad network of regions (Binder et al., 2009). However, questions of how the coordinated network-level activity endows our capacity for flexible conceptual knowledge has been less prominently explored.

Building on the findings from previous chapters, we investigated whether the approach we used to characterize the distributed representation of person-knowledge generalizes to broad conceptual categories. Within regions more active when accessing semantic knowledge (typicality task) compared to the word reading control task, there was little information about conceptual distance both individually (voxel level MVPA) and when we treated these regions as a network (netRSA). This is surprising given the previous research showing sensitivity to semantic distance within these regions at the voxel level (Fairhall & Caramazza, 2013; Liuzzi et al., 2015; Bruffaerts et al., 2013; Devereux, Clarke, Marouchos & Tyler, 2013; Simanova et al., 2012). However, the lack of effect in the netRSA analysis is not inconsistent with previous reports that averaged regional activity within these regions is not modulated by semantic object category. This may indicate that knowledge about semantic categories is represented differently within the brain semantic network from people-knowledge. Another potential explanation for the lack of significance in netRSA analysis is the different cognitive tasks investigated. Whereas during person-knowledge we investigated a single category and manipulate cognitive task, in the last study investigated the effect of semantic class with a single typically of task – which may not influence average regional responses in the same fashion.

Alternatively, this could be because of our voxel selection procedure. Contrasting all categories versus control condition may have imposed a homogenizing effect on the activity patterns – diluting regional specialization. This methodological decision may potentially account for the lack of both netRSA and individual regions RSA results.

When we selected regions that varied the most across categories, we observed a network of regions where differences between the multivariate patterns of activity conformed to our a priori semantic model. A result that indicates that semantic representation of object concepts is encoded across the ensemble activations patterns of this network. Moreover, we directly compared local and distributed information content by comparing the a priori semantic model fit of netRSA to the combined fit of voxel level RDMs in single regions. We observed that, within the regions showing the greatest sensitivity to semantic category, general semantic knowledge is strongly represented in the coordinated activity of these regions. A pattern that we previously saw in the person-knowledge. Taken together, these results show that coordinated activity across regions encodes object category relationships. Overall, similar to the manner in which person-knowledge relies on co-ordination across regions subtly tuned to different cognitions – understanding of semantic distance relies of the interplay of regions tuned to different categories. However, the analytic approaches used seem particularly sensitive to the ROI selection process. While independent ROIs selection (chapters 3 and 4) or ROI selection based on variation across conditions produces sensitive ROIs, selection based against a single referent condition appear to homogenize differential effects. Future work may benefit from focusing on regions defined via voxel-level multivariate pattern analysis (c.f. Fairhall & Caramazza, 2013), which could result in increased sensitivity to semantic distance.

Broader Implications

Mechanisms of semantic representation is a topic of debate. Questions regarding grounded/disembodied representations, hub vs distributed nature of these representations are topics of active debate. Emerging evidence from studies of perception suggest that different forms of information are represented in the coordinated activity across multiple regions. Wang et al. (2016) demonstrated that different stimulus categories such as faces, scenes animals and tools can be discriminated by whole-brain connectivity patterns across different regions. Authors showed that alterations in connectivity between regions carry information about stimulus category being viewed – lending credence to the idea that information can be encoded in the cross-regional communication. The effect persisted even when activation magnitudes in category preferring regions was regressed out as well as when category preferring regions were excluded from the analysis – suggesting that information sufficient for category discrimination is present in the distributed correlation patterns between traditionally category-general regions.

Further stressing the importance of distributed knowledge encoding - Fang et al (2018) showed that damage to white matter tracts results in impairments not explained by the functional properties of the regions they connect. Authors constructed an RDM representing object-naming impairments associated with lesions in different white matter tracts. Using RSA, authors showed that white matter RDMs could not be explained by RDMs in the grey matter regions that were being connected.

Evidence presented in the experimental chapters of this thesis suggest that human conceptual knowledge might be similarly encoded in the interactions across multiple – general purpose regions. These results are seemingly at odds with theories which suggest that ATL as an integration hub of semantic knowledge (Patterson et al., 2007). However, it is consistent with a view which posits ATL as a part of a network of regions exhibiting hub-like behavior (see next section).

Single hub or multiple hubs.

The role of ATL in semantic knowledge has been well established. ATL is active during a variety of semantic tasks (Visser, Jefferies & Lambon-Ralph, 2010), and focal damage to the ATL (such as in semantic dementia SD) results in debilitating semantic deficits. TMS applied to either left or right temporal lobe produces semantic impairments (Pobric, Jefferies & Lambon-Raph. 2010), consistent with the notion that ATLs are part of a semantic network. However, there is an ongoing debate whether it constitutes a sole semantic hub. Can, and if so, how can elaborate semantic knowledge emerge without an explicit hub? Distributed accounts, such as Tyler & Moss (2001) conceptual structure account suggest that robust semantic representation emerges without a conceptual hub.

Binder et al (2009) described a network of regions is active for broad kinds of stimuli – potentially implicating multiple regions in hub-like activity. Pulvermüller (2013) discussed five candidates for a semantic hub, including but not limited to the ATL. Evidence reviewed by Pulvermüller (2013) suggest that in addition to the ATL - IFG,

Superior Temporal Gyrus, AG and pMTG, all receive processed, multi-modal input and have been reported to play a central role in semantic processing.

Semantic Dementia (SD) affects both anterior and posterior temporal lobes, and while ATL is the most common site of atrophy, result do not rule out possible involvement of pMTG (Hickok & Poeppel, 2004). Semantic deficits similar to those of SD are associated with acute left posterior temporal lobe damage (Hickok & Poeppel, 2004), potentially implicating the pMTG as well as ATL in semantic knowledge. Neuroimaging studies suggest that lateral while ATL is involved in accessing knowledge about individual exemplars including people and places (Fairhall et al., 2013) it is relatively uninvolved in retrieving category level information (Fairhall & Caramazza, 2013). Fairhall & Caramazza (2013) showed that during category-level retrieval, pMTG and the precuneus contain rich information about conceptual distance.

Left ATL resection in epileptic patients tends to impaired naming (in some groups), but not comprehension abilities, suggesting that ATL can play crucial a role in naming, but not necessarily overall semantic representation (Ives-Deliperi & Butler, 2012). In line with these findings - Bi et al. (2011) described a glioma patient for whom left ATL resection produced naming, but not conceptual deficits, calling into question ATL role as a sole conceptual hub.

These results, together with evidence presented in this thesis suggest the presence of not a single hub located in the ATL, but a network of regions containing rich conceptual information, jointly encoding conceptual structure. One of the clearest illustrations of this was in Chapter 5, where we present evidence that knowledge about

semantic distance about object concepts is represented in an ensemble of regions, more than in single regions. In Chapter 3 we demonstrated the generality of the semantic system. We presented evidence that regions involved in person-knowledge respond to multiple kinds of semantic demands, and person-knowledge structure is encoded across these regions. Chapter 4 provided evidence that conceptual structure encoded is robust and cannot be explained away by perceptual effects. Together these results further the notion that conceptual knowledge relies on multiple interacting regions.

System redundancy

We observed that the brain semantic system responds to a broad range of stimuli (object concepts of different categories, different kinds of person knowledge, written and pictorial stimuli). For any particular stimulus, multiple brain regions are engaged. This raises a question whether these regions process redundant information.

Other brain networks exhibit similarly broad-tunings. For example, visual attention recruits a network of regions spanning the frontal and parietal lobes. While often recruited together, these regions perform different tasks. Across multiple studies, functional neuroimaging studies delineated a ventral aspect involved in the detection of salient, exogenous, stimuli detection and dorsal attention system responsible for endogenous attention orientation (Corbetta & Shulman, 2002). It is likely that generality in the semantic system reflects something similar – namely that regions involved in conceptual knowledge perform not redundant, but complementary functions.

Two pieces of evidence from this thesis indicate that representation contained within regions of this network are not redundant with one another. The first is the differential response profiles observed across regions shown in stacked bar plots in Chapter 3 (c.f. figure 2.3). With respect to the representation of person related attributes, it can be seen that person-knowledge is not represented in all regions the same way – regions exhibit different tunings. This suggests that activation, while being present for each of the cognitive domains, had differential, non-redundant, weighting within regions of this network. Following on from this, we observe that regional clustering based on responsive profile across different person knowledge tasks (chapters 3 & 4, c.f. figures 2.4, 3.3) groups regions into more similar (redundant) or different (diverse) functional profiles. Clustering regions based on response pattern similarity, in effect, measures redundancy – and would be evident in these plots. Instead, while we observe similarity between left and right hemisphere homologues – brain regions are grouped apart, indicating heterogenic rather than redundant contributions. Likewise, it is this diversity in the regional response profile that allow the taxonomic reconstruction of the relationship between different cognitive processes or between the representation of different categories of knowledge. Collectively these results suggest that brain regions contribute to a higher-order, network-level signature, jointly encoding conceptual knowledge. Just like voxel patterns in an individual region jointly encode stimulus features in a distributed code (Haxby et al, 2001), evidence presented in this thesis suggest that regions contribute to a distributed network level-code. The netRSA procedure we used throughout this thesis discounts local code - by averaging across voxels within a region. In Chapter

5 we demonstrated that conceptual distance between object categories is reflected in locally averaged network code. When we directly compared locally averaged netRSA versus combined voxel activity from all regions of interest, we observed that they capture similar levels of information. If regional response patterns were redundant, we would expect the information contained across the network to not differ from information contained within individual regions. Our results instead suggest the presence of sufficient heterogeneity across regions to facilitate robust encoding of conceptual distance.

Collectively, these results suggest a system where different (but overlapping) forms of knowledge are represented in different brain regions and that these regions work together to share information and endow the flexible combination of semantic information about the world. It is still an open question of *how* these different brain regions work together to allow us our rapid and effortless access to stored knowledge - one mechanism might be the parallel activation of elements of the semantic system.

System Speed

There is evidence for some parallel computation within the semantic system. In an MEG experiment Sudre et al (2012) suggested that low-level perceptual stimulus features can be decoded early (50ms) in the brain, while different semantic features can be decoded later (250ms - 450ms time window). Areas involved in semantic processing activate in parallel or in rapid (near-simultaneous) succession (Pulvermüller et al., 2009). There are subtle differences (10-20ms) between time-courses of broadly distributed semantic regions (STG, IFG), consistent with near-parallel accounts of semantic activation. The

amplitude of N400 meaning component, typically occurring in the 300ms-500ms window is modulated by semantic content, such as face familiarity or semantic anomaly (Debrulle, Pineda & Renault, 1996; Lau, Phillips & Poeppel, 2008). Combined evidence showing that N400 component arises from a distributed network of sources (Federmeier, Kutas & Dickson, 2016) and is modulated in amplitude rather than latency (Federmeier et al., 2016) can be reflective of the coordinated output from a network of regions.

Abstract words, emotion as a feature

There are differences in how abstract and concrete words are encoded, demonstrated by differential impairments in stroke patients (Crutch & Warrington, 2004). However, a recent meta-analysis demonstrates that abstract and concrete words might be represented in the same network of regions (Wang, Conder, Blitzer & Shinkareva, 2010). Authors found that abstract concepts preferentially engage the LATL & LIFG, while concrete concepts preferentially engage PCC, PHG, pMTG and AG. That brain regions involved in concrete nouns processing, contain information about abstract word meaning similarity as well – suggesting at least somewhat overlapping neural substrates for concrete and abstract concepts. Furthermore, recent accounts suggest abstract knowledge is encoded similarly to concrete knowledge, if emotion, sociality are included as features. Abstract words tend to be more emotionally valenced than concrete nouns, and if affect is controlled for, known reaction time advantage for abstract words disappear (Kousta, Vigliocco & Andrews, 2011). Combined these results suggest that while abstract concepts rely on different

features than concrete nouns, the encoding principles might be more similar than previously thought.

Further directions

This thesis focused on coordinated network activity rather than processing happening within individual regions. Further research could benefit from targeted investigations of regional contributions to network function. Incorporating functional connectivity measures can also further inform relationships between region and network function.

Large scale brain networks, like the one involved in person knowledge, is comprised of coordinating functional units. In chapters 3 & 4 We demonstrated that within core and extended systems for person knowledge, further divisions are visible based on regional tuning similarity. We observed that regions associated with internalised cognition (previously known as DMN) contribute shared response similarity across multiple person-knowledge tasks. What exactly these regions contribute is rather unclear. Modern informational connectivity measures make it possible to track how information is passed and transformed between regions, or regional contributions to network-level function. Measures such as combinatorial MVPA (Wang et al., 2016) or modelling regional transformations across the networks (Anzellotti et al., 2017) enable us to track how information is transformed along processing steps. Distilling contributions of regions (or groups of regions) to network-level function would be a step forward in understanding how conceptual knowledge emerges from regional interactions.

Incorporating traditional functional connectivity measures can help uncover precise connections that are important for specific kinds of semantic knowledge encoding. Wang et al. (2016) demonstrated that information meaningful to stimulus category discrimination is present in groups of pairwise regional connectivity patterns. Building on evidence presented in this thesis, namely that groups of high-level, broadly-tuned semantic regions jointly encode information – further insights are possible. For example, which functional connections are most informative about conceptual categories? Are any of them redundant? This line of research would help better understand how information is passed between brain regions during semantic processing.

Current thesis stressed the importance of network function. We presented evidence that robust information is encoded even at a coarse network level activity. Next logical avenue to pursue is investigate the contributions of individual regions to network level code.

Conclusion

Combined, evidence presented in this thesis speak to the distributed nature of semantic representation. Extending accounts that assume a single hub region – we stress the importance of a network of regions jointly encoding conceptual knowledge. Different kinds of person-knowledge and information about semantic distance of object categories emerge as a product of the coordinated activity across multiple regions. For person-knowledge, we observed that most-all regions are involved in most-all person-related cognitive domains but that subtle differences between regional profiles contribute to distributed

representations of knowledge. We observe conceptual knowledge about objects is similarly distributed across the broader semantic system – network activity conforms to semantic distance model – and to a greater extent than the within-region multivariate pattern when combined across regions.

References

- Adolphs R. 2002. Neural systems for recognizing emotion. 12:169–177.
- Adolphs R. 2003. Cognitive neuroscience: Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*. 4:165–178.
- Adolphs R. 2009. The Social Brain: Neural Basis of Social Knowledge. 60:693–716.
- Anderson, A. J., Lalor, E. C., Lin, F., Binder, J. R., Fernandino, L., Humphries, C. J., ... & Wang, X. (2018). Multiple regions of a cortical network commonly encode the meaning of words in multiple grammatical positions of read sentences. *Cerebral Cortex*.
- Anzellotti, S., & Coutanche, M. N. (2018). Beyond functional connectivity: Investigating networks of multivariate representations. *Trends in cognitive sciences*, 22(3), 258-269.
- Anzellotti, S., Caramazza, A., & Saxe, R. (2017). Multivariate pattern connectivity. *bioRxiv*, 046151.
- Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in cognitive sciences*, 7(2), 84-91.
- Beaty, R. E., Benedek, M., Wilkins, R. W., Jauk, E., Fink, A., Silvia, P. J., ... & Neubauer, A. C. (2014). Creativity and the default network: A functional connectivity analysis of the creative brain at rest. *Neuropsychologia*, 64, 92-98.

- Beauchamp, M. S., & Martin, A. (2007). Grounding object concepts in perception and action: evidence from fMRI studies of tools. *Cortex*, 43(3), 461-468.
- Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex*, 10(3), 295-307.
- Bi, Y., Wang, X., & Caramazza, A. (2016). Object domain and modality in the ventral visual pathway. *Trends in cognitive sciences*, 20(4), 282-290.
- Bi, Y., Wei, T., Wu, C., Han, Z., Jiang, T., & Caramazza, A. (2011). The role of the left anterior temporal lobe in language processing revisited: evidence from an individual with ATL resection. *cortex*, 47(5), 575-587.
- Binder JR, Desai RH, Graves WW, Conant LL. 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex*. 19:2767–2796.
- Binder, J. R., & Fernandino, L. (2015). Semantic Processing. *Brain Mapping: An Encyclopedic Reference*
- 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(12), 2767-2796.
- Bortolon C, Capdevielle D, Raffard S. 2015. Face recognition in schizophrenia disorder: A comprehensive review of behavioral, neuroimaging and neurophysiological studies. *Neuroscience & Biobehavioral Reviews*. 53:79–107.

- Bracci, S., Cavina-Pratesi, C., Ietswaart, M., Caramazza, A., & Peelen, M. V. (2011). Closely overlapping responses to tools and hands in left lateral occipitotemporal cortex. *Journal of neurophysiology*, *107*(5), 1443-1456.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British journal of psychology*, *77*(3), 305-327.
- Bruffaerts, R., Dupont, P., Peeters, R., De Deyne, S., Storms, G., & Vandenberghe, R. (2013). Similarity of fMRI activity patterns in left perirhinal cortex reflects semantic similarity between words. *Journal of Neuroscience*, *33*(47), 18597-18607.
- Bzdok D, Langner R, Caspers S, Kurth F, Habel U, Zilles K, Laird A, Eickhoff SB. 2010. ALE meta-analysis on facial judgments of trustworthiness and attractiveness. *Brain Struct Funct*. 215:209–223.
- Calder AJ, Young AW. 2005. Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*. 6:641–651.
- Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of cognitive neuroscience*, *10*(1), 1-34.
- Caramazza, A., Anzellotti, S., Strnad, L., & Lingnau, A. (2014). Embodied cognition and mirror neurons: a critical assessment. *Annual review of neuroscience*, *37*, 1-15.
- Carbon, C. C., Grüter, T., Grüter, M., Weber, J. E., & Lueschow, A. (2010). Dissociation of facial attractiveness and distinctiveness processing in congenital prosopagnosia. *Visual Cognition*, *18*(5), 641-654.

- Chatterjee, A., Thomas, A., Smith, S. E., & Aguirre, G. K. (2009). The neural response to facial attractiveness. *Neuropsychology*, *23*(2), 135.
- Cichy, R. M., Khosla, A., Pantazis, D., & Oliva, A. (2017). Dynamics of scene representations in the human brain revealed by magnetoencephalography and deep neural networks. *Neuroimage*, *153*, 346-358.
- Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature neuroscience*, *17*(3), 455.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, *3*(3), 201.
- Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*(2), 163.
- Crutch, S. J., & Warrington, E. K. (2004). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*(3), 615-627.
- Crutch, S. J., & Warrington, E. K. (2005). Gradients of semantic relatedness and their contrasting explanations in refractory access and storage semantic impairments. *Cognitive Neuropsychology*, *22*(7), 851-876.
- Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., & Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition*, *92*(1-2), 179-229.
- Debruille, J. B., Pineda, J., & Renault, B. (1996). N400-like potentials elicited by faces and knowledge inhibition. *Cognitive Brain Research*, *4*(2), 133-144.

- Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *Journal of Neuroscience*, *33*(48), 18906-18916.
- Devlin, J.T., Russell, R.P., Davis, M.H., Price, C.J., & Tyler, L.K. (2002). Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. *Neuropsychologia*, *40*, 54-75.
- di Oleggio Castello, M. V., Halchenko, Y. O., Guntupalli, J. S., Gors, J. D., & Gobbini, M. I. (2017). The neural representation of personally familiar and unfamiliar faces in the distributed system for face perception. *Scientific reports*, *7*(1), 12237.
- Downing, P. E., & Peelen, M. V. (2011). How might occipitotemporal body-selective regions interact with other brain areas to support person perception?. *Cognitive Neuroscience*, *2*(3-4), 216-226.
- 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.
- Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science*, *1*, 393-416.
- Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.*, *30*, 123-152.
- Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392: 598–601.

- Fairhall SL, Anzellotti S, Ubaldi S, Caramazza A. 2014. Person- and place-selective neural substrates for entity-specific semantic access. *Cereb Cortex*. 24:1687–1696.
- Fairhall SL, Caramazza A. 2013a. Category-selective neural substrates for person- and place-related concepts. *Cortex*. 49:2748–2757.
- Fairhall SL, Caramazza A. 2013b. Brain regions that represent amodal conceptual knowledge. *J Neurosci*. 33:10552–10558.
- Fairhall SL, Ishai A. 2007. Effective Connectivity within the Distributed Cortical Network for Face Perception. 17:2400–2406.
- Fang, Y., Han, Z., Zhong, S., Gong, G., Song, L., Liu, F., ... & He, Y. (2015). The semantic anatomical network: Evidence from healthy and brain-damaged patient populations. *Human brain mapping*, 36(9), 3499-3515
- Fang, Y., Wang, X., Zhong, S., Song, L., Han, Z., Gong, G., & Bi, Y. (2018). Semantic representation in the white matter pathway. *PLoS biology*, 16(4), e2003993.
- Federmeier, K. D., Kutas, M., & Dickson, D. S. (2016). A common neural progression to meaning in about a third of a second. In *Neurobiology of language* (pp. 557-567). Academic Press.
- Federmeier, K. D., Kutas, M., & Dickson, D. S. (2016). A common neural progression to meaning in about a third of a second. In *Neurobiology of language* (pp. 557-567). Academic Press.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *PNAS*. 102:9673–9678.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, 102(27), 9673–9678. <http://doi.org/10.1073/pnas.0504136102>

Frith, C. D. (2007). The social brain?. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362(1480), 671-678.

Gauthier I, Tarr MJ, Moylan J, Skudlarski P. 2000. The fusiform “face area” is part of a network that processes faces at the individual level. *Journal of cognitive ...*

Gerlach, C. (2007). A review of functional imaging studies on category specificity. *Journal of Cognitive Neuroscience*, 19(2), 296-314.

Gesierich B, Jovicich J, Riello M, Adriani M, Monti A, Brentari V, Robinson SD, Wilson SM, Fairhall SL, Gorno-Tempini ML. 2011. Distinct Neural Substrates for Semantic Knowledge and Naming in the Temporoparietal Network. 22:bhr286–bhr2226.

Glosser G, Salvucci AE, Chiaravalloti ND. 2003. Naming and recognizing famous faces in temporal lobe epilepsy. *Neurology*. 61:81–86.

Gobbini MI, Haxby JV. 2007. Neural systems for recognition of familiar faces. *Neuropsychologia*. 45:32–41.

- Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006). Perceptual knowledge retrieval activates sensory brain regions. *Journal of Neuroscience*, *26*(18), 4917-4921.
- Gorno-Tempini ML, Price CJ, Josephs O, Vandenberghe R, Cappa SF, Kapur N, Frackowiak RS, Tempini ML. 1998. The neural systems sustaining face and proper-name processing. *Brain*. 121 (Pt 11):2103–2118.
- Grabowski TJ, Damasio H, Tranel D, Ponto LLB, Hichwa RD, Damasio AR. 2001. A role for left temporal pole in the retrieval of words for unique entities. *Human Brain Mapping*. 13:199–212.
- Greene JD, Sommerville RB, Nystrom LE, Darley JM, Cohen JD. 2001. An fMRI investigation of emotional engagement in moral judgment. *Science*. 293:2105–2108.
- Greer, M. J., Van Casteren, M., McLellan, S. A., Moss, H. E., Rodd, J., Rogers, T., & Tyler, L. K. (2001). The emergence of semantic categories from distributed featural representations. In *Proceedings of the 23rd annual conference of the Cognitive Science Society* (pp. 358-363). London: Lawrence Erlbaum Associates.
- Gutiérrez-García, A., Beltrán, D., & Calvo, M. G. (2018). Facial attractiveness impressions precede trustworthiness inferences: lower detection thresholds and faster decision latencies. *Cognition and Emotion*, 1-8.
- Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in cognitive sciences*, *11*(7), 299-306.

- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41(2), 301-307.
- Haxby JV, Hoffman EA, Gobbini MI. 2000. The distributed human neural system for face perception. *Trends in Cognitive Sciences*. 4:223–233.
- Heide Von Der RJ, Skipper LM, Olson IR. 2013. Anterior temporal face patches: a meta-analysis and empirical study. 7.
- Hein G, Knight RT. 2008. Superior Temporal Sulcus—It's My Area: Or Is It? <http://dxdoiorg/101162/jocn200820148>. 20:2125–2136.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2), 67-99.
- Horowitz, S. G., Braun, A. R., Carr, W. S., Picchioni, D., Balkin, T. J., Fukunaga, M., & Duyn, J. H. (2009). Decoupling of the brain's default mode network during deep sleep. *Proceedings of the National Academy of Sciences*.
- Humphreys, G. F., Hoffman, P., Visser, M., Binney, R. J., & Ralph, M. A. L. (2015). Establishing task-and modality-dependent dissociations between the semantic and default mode networks. *Proceedings of the National Academy of Sciences*, 201422760.
- Huth AG, de Heer WA, Griffiths TL, Theunissen FE, Gallant JL. 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*. 532:453–458.

- Ishai, A., Haxby, J. V., & Ungerleider, L. G. (2002). Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *Neuroimage*, 17(4), 1729-1741.
- Ives-Deliperi, V. L., & Butler, J. T. (2012). Naming outcomes of anterior temporal lobectomy in epilepsy patients: a systematic review of the literature. *Epilepsy & Behavior*, 24(2), 194-198.
- Kan IP, Barsalou LW, Solomon KO, Minor JK, Thompson-Schill SL. 2010. Role of mental imagery in a property verification task: fmri evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology*. 20:525–540.
- Kanwisher N, McDermott J, Chun MM. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*. 17:4302–4311.
- Keidel, J. L., Oedekoven, C. S., Tut, A. C., & Bird, C. M. (2017). Multiscale integration of contextual information during a naturalistic task. *Cerebral Cortex*, 1-9.
- Koster-Hale J, Richardson H, Velez N, Asaba M, Young L, Saxe R. 2017. Mentalizing regions represent distributed, continuous, and abstract dimensions of others' beliefs. *NeuroImage*. 161:9–18.
- 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.

- Kriegeskorte, N. (2009). Relating population-code representations between man, monkey, and computational models. *Frontiers in Neuroscience*, 3, 35.
- Kriegeskorte, N., & Mur, M. (2012). Inverse MDS: Inferring dissimilarity structure from multiple item arrangements. *Frontiers in psychology*, 3, 245.
- Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences*, 104(51), 20600-20605.
- Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 2, 4.
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics:(de)constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920.
- Lindner, M., Hundhammer, T., Ciaramidaro, A., Linden, D. E., & Mussweiler, T. (2008). The neural substrates of person comparison—an fMRI study. *Neuroimage*, 40(2), 963-971.
- Liuzzi, A. G., Bruffaerts, R., Dupont, P., Adamczuk, K., Peeters, R., De Deyne, S., ... & Vandenberghe, R. (2015). Left perirhinal cortex codes for similarity in meaning between written words: comparison with auditory word input. *Neuropsychologia*, 76, 4-16.
- Mahon, B. Z., & Caramazza, A. (2005). The orchestration of the sensory-motor systems: Clues from neuropsychology. *Cognitive neuropsychology*, 22(3-4), 480-494.

- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of physiology-Paris*, 102(1-3), 59-70.
- Martin A, Chao LL. 2001. Semantic memory and the brain: structure and processes. *Current Opinion in Neurobiology*. 11:194–201.
- Martin, A., & Caramazza, A. (2003). Neuropsychological and neuroimaging perspectives on conceptual knowledge: An introduction. *Cognitive Neuropsychology*, 20(3-6), 195-212.
- Minnebusch DA, Daum I. 2009. Neuropsychological mechanisms of visual face and body perception. *Neuroscience & Biobehavioral Reviews*. 33:1133–1144.
- Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K. M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *science*, 320(5880), 1191-1195.
- Moeller S, Freiwald WA, Tsao DY. 2008. Patches with links: a unified system for processing faces in the macaque temporal lobe. *Science*. 320:1355–1359.
- Moran JM, Lee SM, Gabrieli JDE. 2011. Dissociable Neural Systems Supporting Knowledge about Human Character and Appearance in Ourselves and Others. http://dxdoiorg/101162/jocn_a_00009. 23:2222–2230.
- Moss, H. E., Rodd, J. M., Stamatakis, E. A., Bright, P., & Tyler, L. K. (2004). Anteromedial temporal cortex supports fine-grained differentiation among objects. *Cerebral cortex*, 15(5), 616-627.

- Nielson, K. A., Seidenberg, M., Woodard, J. L., Durgerian, S., Zhang, Q., Gross, W. L., ... & Rao, S. M. (2010). Common neural systems associated with the recognition of famous faces and names: an event-related fMRI study. *Brain and cognition, 72*(3), 491-498.
- Noppeney, U., Patterson, K., Tyler, L. K., Moss, H., Stamatakis, E. A., Bright, P., ... & Price, C. J. (2007). Temporal lobe lesions and semantic impairment: a comparison of herpes simplex virus encephalitis and semantic dementia. *Brain, 130*(4), 1138-1147.
- Olson IR, McCoy D, Klobusicky E, Ross LA. 2013. Social cognition and the anterior temporal lobes: a review and theoretical framework. *Soc Cogn Affect Neurosci. 8*:123–133.
- Olson IR, Plotzker A, Ezzyat Y. 2007. The Enigmatic temporal pole: a review of findings on social and emotional processing. *Brain. 130*:1718–1731.
- Papeo, L., Lingnau, A., Agosta, S., Pascual-Leone, A., Battelli, L., & Caramazza, A. (2014). The origin of word-related motor activity. *Cerebral Cortex, 25*(6), 1668-1675.
- Patterson K, Nestor PJ, Rogers TT. 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience. 8*:976–987.
- Patterson, K., & Ralph, M. A. L. (2016). The hub-and-spoke hypothesis of semantic memory. In *Neurobiology of language*(pp. 765-775). Academic Press.

- Pobric, G., Jefferies, E., & Ralph, M. A. L. (2010). Amodal semantic representations depend on both anterior temporal lobes: evidence from repetitive transcranial magnetic stimulation. *Neuropsychologia*, *48*(5), 1336-1342.
- Podgorny, P., & Garner, W. R. (1979). Reaction time as a measure of inter- and intraobject visual similarity: Letters of the alphabet. *Perception & Psychophysics*, *26*(1), 37-52.
- Price CJ. 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*. *62*:816–847.
- Pulvermüller, F. (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends in cognitive sciences*, *17*(9), 458-470.
- Pulvermüller, F., Shtyrov, Y., & Hauk, O. (2009). Understanding in an instant: neurophysiological evidence for mechanistic language circuits in the brain. *Brain and language*, *110*(2), 81-94.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(2), 676-682.
- Rajimehr R, Young JC, Tootell RBH. 2009. An anterior temporal face patch in human cortex, predicted by macaque maps. *PNAS*. *106*:1995–2000.
- Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42.

- Rissman, J., Chow, T. E., Reggente, N., & Wagner, A. D. (2016). Decoding fMRI signatures of real-world autobiographical memory retrieval. *Journal of cognitive neuroscience*, 28(4), 604-620.
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: The MIT Press.
- Rogers, T. T., Ralph, L., Matthew, A., Garrard, P., Bozeat, S., McClelland, J. L., ... & Patterson, K. (2004). Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychological review*, 111(1), 205.
- Ross LA, Olson IR. 2012. What's Unique about Unique Entities? An fMRI Investigation of the Semantics of Famous Faces and Landmarks. 22:2005–2015.
- Saxe R, Powell LJ. 2006. It's the thought that counts: specific brain regions for one component of theory of mind. *Psychol Sci*. 17:692–699.
- Schacter DL, Addis DR. 2007. The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*. 362:773–786.
- Schacter, D. L., & Addis, D. R. (2007). Constructive memory: the ghosts of past and future. *Nature*, 445(7123), 27.
- Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The future of memory: remembering, imagining, and the brain. *Neuron*, 76(4), 677-694.

- Schilbach, L., Eickhoff, S. B., Rotarska-Jagiela, A., Fink, G. R., & Vogeley, K. (2008). Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the “default system” of the brain. *Consciousness and cognition*, 17(2), 457-467.
- Seidenberg, M., Guidotti, L., Nielson, K. A., Woodard, J. L., Durgerian, S., Zhang, Q., ... & Rao, S. M. (2009). Semantic knowledge for famous names in mild cognitive impairment. *Journal of the International Neuropsychological Society*, 15(1), 9-18.
- Semenza C. 2011. Naming with proper names: the left temporal pole theory. *Behavioural Neurology*. 24:277–284.
- Simanova, I., Hagoort, P., Oostenveld, R., & van Gerven, M. (2014). Surface-based searchlight mapping of modality-independent responses to semantic categories using high-resolution fMRI. In *the Sixth Annual Meeting of the Society for the Neurobiology of Language (SNL 2014)*.
- Simanova, I., Hagoort, P., Oostenveld, R., & Van Gerven, M. A. (2012). Modality-independent decoding of semantic information from the human brain. *Cerebral cortex*, 24(2), 426-434.
- Simmons WK, Martin A. 2009. The anterior temporal lobes and the functional architecture of semantic memory. *J Inter Neuropsych Soc*. 15:645.
- Simmons, W. K., Bellgowan, P. S., & Martin, A. (2007). Measuring selectivity in fMRI data. *Nature neuroscience*, 10(1), 4.

- Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia*, *45*(12), 2802-2810.
- Spreng RN, Stevens WD, Chamberlain JP, Gilmore AW, Schacter DL. 2010. Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *NeuroImage*. 53:303–317.
- Spreng, R. N., & Grady, C. L. (2010). Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network. *Journal of cognitive neuroscience*, *22*(6), 1112-1123.
- Spunt RP, Adolphs R. 2017. A new look at domain specificity: insights from social neuroscience. *Nature Reviews Neuroscience*. 18:559–567.
- Sudre, G., Pomerleau, D., Palatucci, M., Wehbe, L., Fyshe, A., Salmelin, R., & Mitchell, T. (2012). Tracking neural coding of perceptual and semantic features of concrete nouns. *NeuroImage*, *62*(1), 451-463.
- Thompson-Schill SL. 2003. Neuroimaging studies of semantic memory: inferring "how" from "where". *Neuropsychologia*. 41:280–292.
- Thompson-Schill, S. L., Kan, I. P., & Oliver, R. T. (2006). Functional neuroimaging of semantic memory. *Handbook of functional neuroimaging of cognition*, *2*, 149-190.
- Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *The quarterly journal of experimental psychology*, *37*(4), 571-590.

- Tippett LJ, Miller LA, Farah MJ. 2000. Prosopamnesia: a selective impairment in face learning. *Cognitive Neuropsychology*. 17:241–255.
- Todorov A, Gobbini MI, Evans KK, Haxby JV. 2007. Spontaneous retrieval of affective person knowledge in face perception. 45:163–173.
- Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in cognitive sciences*, 12(12), 455-460.
- Tulving E. 1972. Episodic and semantic memory. *Organization of memory*. 1, 381-403
- Tyler, L. K., Moss, H. E., Durrant-Peatfield, M. R., & Levy, J. P. (2000). Conceptual structure and the structure of concepts: A distributed account of category-specific deficits. *Brain and language*, 75(2), 195-231.
- Van Overwalle F. 2009. Social cognition and the brain: A meta-analysis. *Human Brain Mapping*. 30:829–858.
- Van Overwalle F. 2011. A dissociation between social mentalizing and general reasoning. *NeuroImage*. 54:1589–1599.
- 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.
- Wagner AD, Paré-Blagojev EJ, Clark J, Poldrack RA. 2001. Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron*.

- Wang, J., Conder, J. A., Blitzer, D. N., & Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human brain mapping, 31*(10), 1459-1468.
- Wang, X., Fang, Y., Cui, Z., Xu, Y., He, Y., Guo, Q., & Bi, Y. (2016). Representing object categories by connections: evidence from a multivariate connectivity pattern classification approach. *Human brain mapping, 37*(10), 3685-3697.
- Wang, X., Peelen, M. V., Han, Z., Caramazza, A., & Bi, Y. (2016). The role of vision in the neural representation of unique entities. *Neuropsychologia, 87*, 144-156.
- 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.
- Warrington, E. K., & McCarthy, R. (1983). Category specific access dysphasia. *Brain, 106*(4), 859-878.
- Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107*(3), 829-853.
- Werner N-S, Kühnel S, Markowitsch HJ. 2013. The Neuroscience of Face Processing and Identification in Eyewitnesses and Offenders. *Front Behav Neurosci. 7*.
- Wieser MJ, Brosch T. 2012. Faces in Context: A Review and Systematization of Contextual Influences on Affective Face Processing. *Frontiers in Psychology. 3*.
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of personality and social psychology, 37*(3), 395.

Winston, J. S., O'Doherty, J., Kilner, J. M., Perrett, D. I., & Dolan, R. J. (2007). Brain systems for assessing facial attractiveness. *Neuropsychologia*, 45(1), 195-206

Appendix 1

ROI coordinates used in chapter 4.

Coordinates provided in MNI space.

ANOVA contrast

Name	Hemisphere	X	Y	Z
Prec.	M	0	-52	24
RSC	R	14	-56	18
RSC	L	-12	-56	16
pMTG	L	-56	-58	-6
pMTG	R	60	-46	2
PHG	L	-30	-40	-12
PHG	R	34	-34	-16
vmPFC	L	-6	44	2
OFC	L	-24	28	-14
OFC	R	24	26	-14
IPL	L	-52	-48	44
IPL	R	46	-62	46
AG	L	-48	-67	27

All > Control ROIs

Name	Hemisphere	X	Y	Z
vmPFC	M	-4	26	-16
dmPFC	L	-10	46	42
Prec	M	-6	-54	14
VTC	R	26	-18	-18
VTC	L	-34	-30	-20
OFC	L	-40	30	-14
OFC	R	24	28	-12
IFG	L	-48	28	4
IFG	R	48	36	18

pMTG	L	-60	-48	-10
AG	L	-38	-62	30
AG	R	38	-68	34
ATL	L	-62	-16	-20
ATL	R	60	-12	-24