Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

________________________________ _______________
Jonathan H. Drucker Date

Neural Bases of Core and Conceptual Self: Implications for the Representation of Other Persons and Groups of People

By
Jonathan Harris Drucker
Doctor of Philosophy

Psychology

Lawrence W. Barsalou

Advisor

Lynne Nygaard, Ph.D.	James K. Rilling, Ph.D.
Committee Member	Committee Member
Hillary R. Rodman, Ph.D.	Phillip Wolff, Ph.D.
Committee Member	Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.

Dean of the James T. Laney School of Graduate Studies

Date

Neural Bases of Core and Conceptual Self: Implications for the Representation of Other Persons and Groups of People

By

Jonathan Harris Drucker
B.S., Stanford University, 2008
M.A., Emory University, 2011

Advisor: Lawrence W. Barsalou, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2015

Abstract

Neural Bases of Core and Conceptual Self: Implications for the Representation of Other Persons and Groups of People

By Jonathan Harris Drucker

Self-representation is multifaceted. Building on prior research, the current experiment explores two such facets, core and conceptual self, and how these constructs extend to neurocognitive representations of other persons and groups of people. Core self is a representation of oneself as an individual whose subjective experience is unified within a particular moment. It includes identification with one's body (body ownership), authorship for one's actions (agency), and an embodied, egocentric point of view (first-person perspective). Conceptual self is a representation of oneself as an individual whose identity persists through time. It includes one's personality traits and physical characteristics, and the narrative of one's life constructed from accumulated autobiographical memories.

I conducted an fMRI experiment (N = 19) to determine the neural correlates of core and conceptual representations for the self, other persons, and groups of people. On each trial, participants were presented for three seconds with an individual (the self, a relative, a friend, or an acquaintance), a group (adults or children), or a semantic prompt (physical or genetic). Participants were then presented for three seconds with a property (e.g., *tall*) and rated how well the property applied to the individual (e.g., how *tall* is the individual?), group (e.g., how *tall* are children?), or prompt (e.g., to what extent is *tall* a genetic property?). The first phase, in which the individual, group, or prompt was presented, was intended to elicit core representations. The second phase, in which the property was presented, was intended to elicit conceptual representations.

The core self condition recruited brain areas associated with body ownership, agency, the first-person perspective, and visuospatial imagery. The conceptual self condition involved these as well, and further implicated brain areas associated with representing personality traits, semantic person knowledge, and executive control of memory retrieval, decision making, and theory-of-mind. Representations for other persons and groups of people rely on these and other systems insofar as the information they provide is relevant to the task at hand, and available with respect to that person or group of people. Some differences between self and other were modulated by the closeness of the personal relationship.

Neural Bases of Core and Conceptual Self: Implications for the Representation of Other Persons and Groups of People

By

Jonathan Harris Drucker
B.S., Stanford University, 2008
M.A., Emory University, 2011

Advisor: Lawrence W. Barsalou, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2015

Acknowledgments

Although I designed, implemented, and analyzed the experiment in close collaboration with my advisor, Larry Barsalou, the written document is largely my own, especially how I interpreted the results in the Discussion section. Larry would have assessed the results and their implications for the initial hypotheses in a very different manner. The conclusions reached in the Discussion section, together with the arguments for them, are completely my own.

The experimental work presented here was supported by an NIH Director's Pioneer Award DPI OD003312 to Lisa Feldman Barrett at Northeastern University, with a sub-contract to Lawrence Barsalou at Emory University.

The author would also like to thank his parents, Susan Leibmann, M.D. and Mitchell David Drucker, M.D., for their unwavering, unconditional, and selfless support and encouragement.

Table of Contents

Page

List of Tables

List of Figures

List of Supplementary Tables

List of Supplementary Figures

List of Appendices

Introduction

In a previous experiment, I used functional magnetic resonance imaging (fMRI) to explore how the brain instantiates core and conceptual representations of the self, other people, and non-human objects. Those results indicated that core representations differ from conceptual representations in important ways: conceptual representations of humans - though not of objects - recruit a classic network for self-related processing, whereas core representations involve a new set of brain regions that has yet to be studied as a single system. Interestingly, many of the differences observed between representations of different individuals were a matter of degree, rather than categorical. Moving forward here, I further explored these ideas by addressing three issues raised by my previous results. First, does the self enjoy a privileged status in the human brain, or is it a special case of more general processes? Second, do core representations exist for groups of people? If so, are they instantiated by the same network as core representations for individuals? Finally, does activity in the classic self network reflect the predication of properties upon individuals, or the representation of those properties per se? I conducted an fMRI experiment in which participants performed three subtly different tasks designed to address the above issues.

Defining the Self

I define a *person representation* as information in the brain that bears a correspondence with the enduring traits or transient states of an individual person or group of people, whether those traits or states are real or perceived, physical or mental. This definition is based on Rupert's (2011) definition of a representation as a cognitive construct whose correspondence with its referent constitutes knowledge that is useful in navigating one's social and physical world. Self-representation is a type of person representation that pertains specifically to oneself. I will explore two types of self-representation (from now on referred to simply as the *self*), as well as of other people and groups of people, considering the ways in which these various types of representations are overlapping or dissociable.

The self as a construct is relevant across multiple disciplines, and therefore can be difficult to define (Legrand & Ruby, 2009). The *self-concept* is a common avenue of study, defined by Gecas (1982) as "an organization (structure) of various identities and attributes, and their evaluations, developed out of the individual's reflexive, social, and symbolic activities". Neisser (1988), however, noted that the self, treated as a single entity, is full of contradictions: "it is simultaneously physical and mental, public and private, directly perceived and incorrectly imagined, universal and culture-specific". Accordingly, he and future researchers have acknowledged that the self is multifaceted: there is more to the self than just the self-concept. These various interacting components may fruitfully be corralled under the label of *self*, but each may have its own unique characteristics (Gallagher, 2013; Klein & Gangi, 2010; but see Baumeister, 2010).

Two paramount qualities of the self serve to motivate the current experiment. First, the self is unified in a given moment. This is called *synchronicity*, and refers to the fact that a person's current state is attributable to a single individual: *I am currently hungry, a bit warm,* and *standing on my own two feet*. Second, the self is unified across time. This is called *diachronicity*, and refers to the fact that a person's present self represents the same individual as his or her past and future selves: *I was born in Philadelphia, have brown hair,* and *will turn twenty-nine in a few months* (Klein & Gangi, 2010).

Several authors have described the multiplicity of self with considerable sophistication, and their parcellations largely follow the synchronic/diachronic distinction just described. As is often the case in psychology, William James was among the first. He wrote of the "*I*", the subjective self consciously experiencing the world one synchronic moment at a time, willfully interacting with it and pondering one's place in it. By contrast, James's "*Me"* is the diachronic self whose identifiable traits and features are the object of explicit reflection (James, 1892; Tagini & Raffone, 2010). Neisser expands upon these ideas with his five kinds of self (1988). The *ecological* and *interpersonal selves* are conscious agents that interact with the material world and other people, respectively. They are directly perceived, requiring no cognitive reflection; as such, they are synchronic representations of a self unified in a given moment, experienced on a moment-to-moment basis. Neisser's *extended self* is the memory of one's past and anticipation of one's future, whereas the *conceptual self* (i.e., self-concept) represents one's social roles and enduring traits. The *private self* is the result of recognizing that certain elements of one's experience, particularly private thoughts, are not accessible to others. Together, the *extended*, *conceptual*, and *private selves* are diachronic representations of oneself as an enduring entity unified across time. Similarly, Damasio's *core* and *extended selves*, and Gallagher's *minimal* and *narrative selves* obey the same distinction (Damasio, 1999, 2010; Gallagher, 2009).

The current experiment is concerned with the neural correlates of two broad types of self-representation: *conceptual self* and *core self*. I now describe each in turn.

Conceptual Self

What I call *conceptual self* refers to the diachronic representation of oneself as an individual whose identity persists throughout time. Conceptual self can be described from several different angles, but I will focus on two: those that Neisser calls "extended self" and "self-concept". Crucially, all aspects of the conceptual self are reflective: they require a meta-awareness of oneself as a physical and mental being with an identity that persists through time (Baumeister, 2010; Esslen et al., 2008; Gecas 1982).

Memories of the past and anticipations for the future share a common protagonist; it is this common denominator transcending the present, and the recognition that one's identity persists throughout, which form the basis for what has been called the "extended self" or "autobiographical self" (Neisser, 1988; Damasio, 1999; 2010). All of the moments of one's life strung together form a more or less cohesive narrative. The common theme, or what Dennett calls the "center of narrative gravity", is the ever-present protagonist. Thus the conceptual self has also been called the "narrative self" (Dennett, 1993; Gallagher, 2000).

The facts of one's life, gleaned from a collection of individual moments, may be abstracted away from the specific event memories that produced them (Klein & Gangi, 2010). For example, I know that *I loved math as an elementary school student*, even though I cannot remember any particular classes I enjoyed. Similarly, I have no particular memory of my birth in Philadelphia. This more abstract component of the conceptual self, that is, knowledge that is not tied to a particular set of events, is called the self-concept: the representation of oneself as "a particular person in a familiar world" (Neisser, 1988). The self-concept comprises a wide array of knowledge including one's identities (e.g., social roles, religious and ethnic identities, and gender) and evaluations of one's competence and morality (Gecas, 1982). The self-concept arises from a number of different sources, as individuals make inferences about their personal traits from observing their own actions, feelings, and interactions with others, both in the present and in past experiences recalled from memory (Wagner, Haxby, & Heatherton, 2012). One's body of conceptual self-knowledge is far more vast than can be held in mind at any one time: various subsets of the self-concept may be accessed as the need arises. For example, knowledge about one's athletic abilities is more relevant, and therefore more readily accessible, when playing basketball than when taking a math test, when an awareness of one's analytical acumen is more relevant (Markus & Wurf, 1987).

The information contained in the conceptual self is hierarchically organized. For example, conceptual self-knowledge may be organized by level of abstraction, from concrete behaviors (e.g., *I am good at catching baseballs*) to abstract traits (e.g., *I am good at sports*). The structure of the hierarchy is a matter of debate: it is likely that multiple hierarchies coexist, and there may often be violations of strict hierarchical structure; the important thing is that conceptual knowledge is of such an explicit nature that it *can* be organized at multiple levels of abstraction (c.f., Harter, 1983).

Neural substrates of the conceptual self. The typical neuroimaging paradigm for studying the self deals with *self-referential processing*: the cognitive processing of stimuli as they pertain to the self (Northoff, Qin, & Feinberg, 2011). For example, a seminal fMRI experiment (Kelley et al., 2002) presented participants with trait adjectives and asked one of three questions: whether the adjectives described the participant (*self*), a famous person (*other*), or whether the word was printed in uppercase letters (the authors called this control the *semantic* condition, but note that the task did not pertain to semantic meaning). The stimuli were the same for all participants and conditions. The underlying logic is that comparisons between these conditions would eliminate effects specific to the stimuli themselves, but be sensitive to processing of the self (by comparing *self* with *other*) and to processing of person representations in general (by comparing *self* or *other* with the *semantic* control condition). This experiment, and a host of other fMRI and PET studies employing similar logic, have revealed a network for self-related processing (reviews: Northoff & Bermpohl, 2004; Northoff et al., 2006; Qin & Northoff, 2011; van Overwalle, 2009; Wagner et al., 2012). These regions, to which I will refer as the *self network* for the sake of simplicity, are consistently and bilaterally implicated: the ventromedial prefrontal cortex (vmPFC), the dorsomedial prefrontal cortex (dmPFC), and the posterior cingulate cortex (PCC). Less consistently, the precuneus (posterior and superior to the PCC), and the temporoparietal junction (TPJ) have been implicated as well (Legrand & Ruby, 2009). I will explore the relative contributions of these areas to self and other person representation in the discussion section. Interestingly, the self network is conspicuously co-extensive with the default-mode network (DMN), a set of brain regions preferentially activated when a person is at rest and deactivated when a person's attention is directed towards an external task (Buckner et al., 2008; Moran et al., 2013). Stimulus-independent thought (i.e., mind-wandering) often pertains to the self, and this self-referential thought contributes to DMN activation at rest (Schneider et al., 2008; Whitfield-Gabrieli et al., 2011;).

There are two important limitations in the approach just described. First, the self-referential processing paradigm has revealed a neural network for processing the self independent of stimulus modality, but it does not distinguish between the various types of self defined above (Northoff et al., 2006). Second, contrasts between self- and other-related processing reveal that certain areas are preferentially involved in the one or the other, but this does not exclude the possibility that the same areas underlie person representation in general, albeit to varying degrees for different individuals (Denny et al., 2012; Wagner et al., 2012). Taking this a step further, Legrand and Ruby (2009) argue that activation in the canonical self network is not directly related to person representation at all. Rather, they claim that the network is doing something more domain-general: carrying out inferential processes using information activated during memory retrieval. These processes utilize person representation, but are not synonymous with it, and are not specific to the self. Legrand and Ruby go on to suggest that the conceptual self may not be dissociable from conceptual representations of other people, and that future neuroimaging studies should focus on the core self instead. I do not share their pessimism, but their discussion of different types of self-processing was a primary motivation for the current line of research.

Core Self

What I call *core self* refers to the synchronic representation of the self as an individual whose subjective experience is unified within a particular moment (Damasio, 1999; 2010). The core self and the phenomenon of conscious experience are thus intimately related, although the nature of this relationship lies beyond the scope of this paper. Indeed, core self is so grounded in subjective experience that several treatments of the subject are phenomenological in nature, rather than strictly cognitive or neurological (Blanke, 2012; Blanke & Metzinger, 2009; Damasio, 1999). Core self can also be described from several different angles, but I will focus on two: *body-ownership* and *agency*. Crucially, all aspects of the core self are pre-reflective: they are represented independently of the conceptual self-representations described above (Esslen et al., 2008). The core self is experienced directly, and does not require explicit cognitive reflection, unlike the conceptual self (Blanke & Metzinger, 2009).

Body-ownership is the sense that one's body and the parts thereof belong to oneself (Legrand, 2006). It is thought to depend on sensorimotor integration across modalities. Together, visual, tactile, and proprioceptive input provide spatially and temporally concordant information, engendering the sense that they collectively constitute the same self (Jeannerod, 2004). Somatoparaphrenia, the delusion that part of one's body is not one's own, is associated with damage to brain regions involved in multisensory integration, such as the right temporoparietal junction and the insula (Blanke, 2012; Fotopoulou et al., 2010). Somatoparaphrenia can occur in the absence of any perceptual deficits, implying that body-ownership occurs at an association level, transcending individual input modalities (Vallar & Ronchi, 2009).

Agency is the sense that one's actions are the consequence of one's intentions (Nahab et al., 2011). By identifying one's actions as her own, one is able to distinguish herself from the environments with which she constantly interacts; this is the essence of Neisser's ecological self (1988; Haggard & Chambon, 2012). Analogous to how body-ownership is identification with the body, agency is identification with intentional actions. In fact, body-ownership of the effector (i.e., the body part used to carry out an action) may be a prerequisite for a sense of agency for an action (Nahab et al., 2011), although some neural evidence suggests that body-ownership and agency are independent (Tsakiris, Longo, & Haggard, 2010).

Body-ownership and agency require an implicit self/non-self distinction in perception and action planning. "Self-specifying" processes are those that impose this distinction, namely, that one's body and one's actions are under one's own control, and that other objects and events in the environment are not (Christoff, Cosmelli, Legrand, & Thompson, 2011). In this way, self-specifying processes create the synchronic core self in each moment: I am *this*, but I am not *that*. In Neisser's (1988) terminology, the "ecological self" is the core self interacting with its non-self environment, whereas the "interpersonal self" is the core self interacting with non-self people.

The self-specifying process underlying agency is thought to involve a comparison between the predicted and actual consequences of an intentional action. According to the “forward model” of agency, intentional actions (e.g., *moving my arm*) generate sensory predictions (*I will see and feel my arm moving*). When actual sensory input matches these predictions, a sense of agency results, contributing to a sense of self that is unified in the present moment (Blakemore & Frith, 2003). When sensory input does not match any forward predictions, such perceptions are understood to be part of the environment or representative of other individuals’ intentions (Christoff et al., 2011; Gallagher, 2000; Jeannerod, 2004; Legrand & Ruby, 2009).

Neural substrates of the core self. The self-specifying processes that give rise to the core self are associative, requiring the integration of multiple streams of information. Therefore, neuroimaging paradigms for studying body-ownership and agency must manipulate the self-specification process while controlling for low-level stimulus features. One way to manipulate body-ownership is via the rubber hand illusion (RHI). Participants see a fake hand stroked (e.g., with a brush or a cotton swab). At the same time but hidden from view, their own hand is stroked in a similar fashion. If conditions are right (e.g., stroking is synchronous, the real and fake hand are similarly oriented and sufficiently near one another), participants report feeling an illusory ownership of the rubber hand (Blanke, 2012; Ehrsson, 2004; Ehrsson et al., 2005; Tsakiris, 2010). In other words, the rubber hand has been falsely incorporated into their core selves.

Neural activity associated with the RHI and similar illusions (typically contrasted with a perceptually similar control condition without illusory ownership, or with the pre-illusion period) has been demonstrated in the premotor cortex, intraparietal sulcus, temporoparietal junction, posterior insula and frontal operculum, and cerebellum using fMRI (Ehrsson, 2005; Ehrsson et al., 2004; Tsakiris, 2010) and PET (Tsakiris, Hesse, Boy, Haggard, & Fink, 2007).

The sense of agency can be studied by manipulating the temporal latency between an intentional action and sensory feedback of its consequences. One experiment used a finger-tapping paradigm in which visual feedback was provided by a video stream that was either in real-time or delayed by 500 milliseconds. The former condition, in which participants experienced agency for the finger-tapping action, activated the SMA and pre-SMA (jointly referred to as the supplementary motor complex, or SMC), as well as the premotor cortex and cerebellum. Although the SMC activations were the largest found in terms of spatial extent, the authors reported a long list of activation clusters (Tsakiris, Longo, & Haggard, 2010). The SMA is also the source of the readiness potential, an ERP component that precedes intentional actions (Blakemore & Frith, 2003). Awareness of agency for one's actions also activates the right TPJ (Farrer et al., 2008; Nahab et al., 2011), as does a violation of the expectation of agency (Haggard & Chambon, 2012). Apraxia, an inability to perform agentic movements despite intact motor circuitry, has been associated with damage to the left parietal cortex from the IPS all the way down into the TPJ (Gross & Grossman, 2008; Sathian et al., 2011).

Comparing the conceptual and core selves

A few empirical studies have directly compared the neural correlates of the conceptual and core selves. In one fMRI experiment, participants were presented with trait words and instructed to adopt one of two perspectives: the "experiential focus" perspective required participants to monitor their moment-to-moment experience of the adjectives (i.e., core self), whereas the "narrative focus" perspective required participants to reflect on what the adjective meant about them as a person (i.e., conceptual self). Whereas experiential focus elicited activation in the right dorsolateral and ventrolateral prefrontal cortex, secondary somatosensory cortex, and inferior parietal lobule, narrative focus elicited activation in the dorsomedial and ventromedial prefrontal cortex and the posterior cingulate cortex (Farb et al., 2007).

Another experiment compared agentic (i.e., core) and conceptual aspects of the self using fMRI. To elicit the agentic self, participants either selected a card (*self* condition) or watched as one was selected for them (*other* condition). To elicit the conceptual self, participants judged how well a particular adjective applied to themselves or to another person. In both cases, activation for the *self* condition was compared against activation for the *other* condition. Activation related to the agentic self was found in bilateral intraparietal sulcus and right dorsolateral prefrontal cortex, areas involved in goal-directed action planning. Activation related to the conceptual self was found in canonical self areas such as bilateral ventromedial prefrontal cortex, temporoparietal junction, and posterior cingulate cortex, as well as other areas (Powell, Macrae, Cloutier, Metcalfe, & Mitchell, 2010).

A third experiment had participants read three-word sentences about themselves or other people, such as *I am tall* or *She is athletic*, one word at a time while undergoing electroencephalography. The authors surmised that brain activity while participants read the word "I" would correspond with a pre-reflective self (i.e., core self) and that activity while participants read the subsequent adjective would correspond with a reflective self (i.e., conceptual self). During the pre-reflective condition (self > other), left hemisphere activity was found in the insula, medial prefrontal cortex, ventral premotor cortex, fusiform gyrus, and lateral orbitofrontal cortex. Right hemisphere activity was found in the dorsal premotor cortex. During the reflective condition (self > other), left hemisphere activity was found in in the medial prefrontal cortex and lateral temporal cortex. Right hemisphere activity was found in motor and premotor cortices, insula, and medial prefrontal cortex. These results support the idea of a core self and a conceptual self working together, with the balance shifting from the former to the latter, albeit with a great deal of overlap (Esslen et al., 2008). Later, it will be seen that these findings bear a fascinating correspondence with my own results.

By acknowledging the presence of pre-reflective, synchronic aspects of the self, these studies reveal that there is more to the story of self- representation in the brain than the canonical self network. Although they do not establish definitively that core and conceptual self-representations are fully dissociable, these studies indicate that at least somewhat different brain networks underlie processing different types of self.

Comparing self and other

The relationship between representations of oneself and other people is an unresolved puzzle in cognitive neuroscience. One possibility is that they rely on dissociable mechanisms. Much of prior neuroimaging work on the self has operated under this assumption, and it is largely by comparing self- and other-related activity that the regions comprising the self network have been identified (Northoff & Bermpohl, 2004; Northoff et al., 2006; Qin & Northoff, 2011; Wagner et al., 2012). However, many studies have found the opposite effect in these areas, that is, greater activation for other than for self (Gillihan & Farah, 2005; Legrand & Ruby, 2009).

A second possibility, that would explain the conflicting findings, is that the activations found do not reflect person representation at all, but are rather involved in more domain-general cognitive processes engaged by the experimental task. Legrand and Ruby (2009), for example, posit that the forced-choice verification tasks used in these studies require inferential processing based on memory retrieval, and that these "evaluative" processes are driving the observed neural activation. As mentioned earlier, they ultimately suggest that the cognitive neuroscience of the self should abandon conceptually-driven neuroimaging tasks altogether in favor of sensorimotor integration. In other words, the core self should be the focus of future investigation because conceptual processing in the brain is not specific to the self.

A third possibility, which I favor, is that a single system exists for representing the self and other people, and that differences in neural activation reflect a gradient rather than a dichotomy (Northoff et al., 2006; Moran, Kelley, & Heatherton, 2013). In fact it has been suggested that knowledge about the self informs knowledge about other people by serving as a helpful analogy (Decety &

Sommerville, 2003). There is considerable evidence that representations of the self and other people share neurocognitive resources but utilize them differently (Colzato, Wildenberg, & Hommel, 2013; Mashek, Aron, & Boncimino, 2003). For example, the attributions people make about others are influenced by the schemas they use to understand themselves (Markus, Smith, & Moreland, 1985). Similarly, asymmetries in similarity judgments indicate that people use themselves as a reference point when making such judgments (Catrambone, Beike, & Neidenthal, 1996). The memory advantage for stimuli processed with respect to the self, called the *self-reference effect*, is also present for stimuli processed with respect to close, but not distant, others (Wagner et al., 2012). With regard to shared neural resources, the ventromedial prefrontal cortex has been implicated in representing traits of both the self and of a close other, but not of a distant other. Closeness, but not familiarity or similarity, modulated the neural activation in question (Krienen, Tu, & Buckner, 2010; Moran, Lee, & Gabrieli, 2010). Other studies including three recent meta-analyses have found a spatial gradient in the medial prefrontal cortex, such that the closer a person is to the self, the more ventral are the peaks of activation that person is represented (Denny et al., 2012; Mitchell, Macrae, & Banaji, 2006; Murray, Schaer, & Debbané, 2012; van der Meer, Costafreda, Aleman, & David, 2010). Along these lines, I will later talk about “self-relatedness”, a term I use to refer to how close an individual is to the self. For example, oneself or a close friend or relative would be maximally self-related, a stranger would be moderately self-related, and an individual who is not even human would be minimally self-related.

The above evidence supports the idea of a conceptual *other* system that overlaps with the conceptual *self* system. It stands to reason that a *core other* system could exist as well. A case study of a congenital amputee sheds some light on the issue. She was found to understand another person's actions by engaging her own sensorimotor brain systems, but only when she would have been capable of performing the same actions herself. Otherwise, she engaged the medial prefrontal cortex, posterior cingulate/precuneus, and bilateral temporoparietal junction. In other words, she drew upon her core self to understand another person, but only to the extent that such core self knowledge was available. Where it was not, she drew upon more abstract conceptual knowledge instead (Aziz-Zadeh, Sheng, Liew, & Damasio, 2012). A meta-analysis of over 200 fMRI studies came to a similar conclusion. Conditions involving abstract social inferences activated a "mentalizing system" consisting of the medial prefrontal cortex, precuneus, and temporoparietal junction, whereas conditions involving understanding the intentional actions of others recruited the premotor cortex, anterior intraparietal sulcus, and posterior superior temporal sulcus (van Overwalle & Baetens, 2009). These latter regions are analogous to the mirror neuron system, discovered in primates and recently confirmed in humans (Keysers & Gazzola, 2010). Mirror neurons fire both when the organism performs or observes an intentional goal-directed action, suggesting that core self is used as a template for understanding others in this system. It should be noted, however, that controversy remains over whether activation in mirror systems reflects action understanding. Heyes (2010) and Hickok (2013) suggest that mirror activation in humans reflects a learned temporal association between goal-directed actions and sensory feedback related to those actions, but that this does not constitute understanding per se. In this view, information from the mirror system must be integrated at a higher level for genuine understanding to occur. If this is correct, then core representations of others should recruit additional neural regions outside the mirror system.

Prior experiment - overview

In a prior experiment, I investigated the nature of core and conceptual representations for the self, a famous person, and an inanimate object (Drucker et al., submitted). My purpose was twofold. First, I aimed to differentiate between the core self and the conceptual self as they are instantiated in the brain. Second, I aimed to determine whether the core/conceptual distinction could be applied to representations of other people and objects, and to what extent self-representations resemble representations of individuals in general.

The experiment consisted of two tasks conducted in an fMRI scanner. In the *main task*, participants were presented with an *individual* (i.e., a person or thing) and a *property* (Figure 2). Individuals were displayed as one printed word (i.e., “I”, “Oprah”, or “Flame”) on a computer screen, and participants were asked to rate how well the property applied to the individual. The three possible responses were “applies poorly”, “applies somewhat”, and “applies well”. By requiring a judgment of relevance from the participant, I ensured that the participant engaged in deep conceptual processing, as opposed to shallow perceptual processing. On some trials, no property was given. These catch trials allowed the linear separation of the hemodynamic response to the individuals from the hemodynamic response to the properties (Ollinger et al., 2001a; 2001b).

I used three individuals as stimuli: “I” (the self), “Oprah” (another person), and “Flame” (a non-human object that will be described in a moment). “I” represented the self, that is, the participant. On being presented with “I”, participants were asked to think about themselves in preparation for the presentation of the accompanying property. I hypothesized that “I” served to activate the core self, and that the subsequent property and verification task activated the conceptual self, replete with conceptual content.

“Oprah” represented the famous pop culture personality Oprah Winfrey. Prior to scanning, participants read a roughly 300-word biography about her. When presented with “Oprah”, participants were asked to think about the real person Oprah Winfrey in preparation for the presentation of the accompanying property. Thus, “Oprah” served to engender a core representation (should one exist), and the subsequent property phase served to engender a conceptual representation, of a person who is not oneself (i.e., the participant).

Finally, “Flame” represented an historical racecar known as the Blue Flame. Prior to scanning, participants read a roughly 300-word history about the Blue Flame. When presented with “Flame”, participants were asked to think deeply about the actual car in preparation for the accompanying property. Thus, “Flame” served to engender a core representation, and the subsequent property phase served to engender a conceptual representation, of an individual that is a non-human object unrelated to the self (i.e., unlike one’s own car.

A second task served as the baseline for the fMRI data analysis. I opted not to use fixation as a baseline because BOLD activation in many brain regions crucial to self-processing remains high under resting conditions (i.e., the DMN; Northoff & Bermpohl, 2004; Raichle et al., 2001). Using the resting state as a baseline would therefore mask activations in important neural structures, as has been seen previously in the neuroimaging of the self (Mitchell, Heatherton, & Macrae, 2002). Instead, I used an active *baseline task*. In this task, participants simply responded with a button press when they detected a short tone during fixation between trials of the main task. Because this task was active and auditory in nature, as opposed to passive and self-oriented, I anticipated a wide disparity between the BOLD response during this task and the BOLD response during the main task, and consequently a high functional signal-to-noise ratio.

Prior experiment - results

Four analyses yielded interesting results. First, I asked if core representations have anything in common across individuals. To see if a common network subserved core representations for the self (I), another person (Oprah), and a non-human object (Flame), I performed a conjunction of each of the three *individuals* conditions against baseline (Figure 3). Overlap between individuals was substantial: the statistical maps of each against baseline differed mostly in the extent and intensity of the clusters, not in their locations (Figure 4). The set of regions revealed in this analysis (henceforth called the *core self network*) did not overlap with the classic self network (Northoff et al., 2006) not the DMN, implying that these two networks do not play a role in the core representation of individuals.

Second, to see if a common network subserved conceptual representations for the self, another person, and a non-human object, I performed a conjunction of each of the three *properties* conditions (i.e., properties applied to I, Oprah, and Flame) against baseline (Figure 5). The core self network just described for the individuals subsequently remained active for all three *property* conditions, becoming active to an even greater degree than during the *individuals* conditions. Also, the classic self network was activated for both humans (properties for “I” and “Oprah”) but not for the object (“Flame”; see Figure 5). I concluded that the classic self network is involved in instantiating conceptual representations of humans, but not of objects.

Third, I aimed to reveal systematic differences in how the brain instantiates core representations for individuals. I defined a variable called *self-relatedness*, reflecting similarity to the self, with a simple ordinal assignment (“Flame” = 1, “Oprah” = 2, “I” = 3). Using self-relatedness as a linear regressor on activation in the three *individuals* conditions, I found significant clusters immediately adjacent to but largely non-overlapping with the core self network (Figure 6). Additionally, I found large bilateral clusters in the extrastriate body area (Figure 7). All correlations were positive except in the posterior occipital cortex, where they were negative. I suggested that the core self network is indeed recruited in generating core representations for all individuals, but also that a gradient is in effect. Core representations of more self-related individuals are more deeply rooted in body representation and agency (premotor, supplementary motor, and extrastriate body areas), whereas core representations of less self-related individuals are more deeply rooted in visual representations (posterior occipital cortex).

Fourth, I aimed to reveal systematic differences in the way the brain instantiates conceptual representations for individuals. Using the same self-relatedness regressor on activation in the *properties* conditions, I found significant clusters of positive correlation throughout the core self network (that I had identified) and the classic self network (Figure 8). I suggested that the classic self network might be a system that is preferentially, but not exclusively, recruited by the self. In other words, the self/nonself distinction in the brain is graded rather than dichotomous.

Outstanding Issues

The results from my previous experiment raised several interesting follow-up questions that warranted exploration in the current experiment. I will discuss each of these issues in turn.

Issue #1: Does the self enjoy a privileged status in the human brain, or is it a special case of more general processes? One popular assumption is that particular brain networks are dedicated to self-processing: self-processing is, to that extent, modular (for reviews, see Northoff & Bermpohl, 2004; Northoff et al., 2006; but cf. Gillihan & Farah, 2005; Legrand & Ruby, 2009). A common approach based on this assumption is to have participants undergo scanning using PET or fMRI while engaged in tasks (e.g., recalling memories, generating personality traits, or recounting general biographical knowledge) or experiencing stimuli (e.g., faces, bodies, names, actions, attitudes, or traits) that are either self-related or not. Contrasting self with non-self conditions reveals a strikingly consistent network of regions in the medial frontal and parietal cortices, as well as the lateral temporal cortices and temporoparietal junctions. Again, this network overlaps extensively with the DMN, for potential reasons I discussed earlier. One conclusion is that these brain regions constitute a system that underlies the self (Northoff et al., 2006).

The *self-reference effect* provided initial behavioral support for a privileged psychological self. Encoding stimuli in a self-referential way increased the likelihood of subsequent recall (Rogers, Kuiper, & Kirker, 1977). The authors inferred the existence of a “schema” for the self whose activation facilitates the encoding of memories. If a unique self-schema exists, then it stands to reason that it has a neural substrate, as suggested above.

However, although the self-reference effect was replicated many times (Symons & Johnson, 1997), multiple confounding factors have since been discovered, calling into question Rogers and colleagues' conclusions. For example, similar effects of memory enhancement were found for tasks that referred to close others, required evaluative judgments, required complex organization of information, or involved the formation of associations between targets and material in memory (Gillihan & Farah, 2005).

Alternatively, the neural self may emerge from a more domain general set of functions: social and other cognitive processes that are perhaps recruited to a greater degree by self-processing, but are not specific to it. Results from my prior experiment provide some evidence that the degree of activity in relevant brain networks is modulated by *self-relatedness*: how much one personally associates or relates to a stimulus (Northoff et al., 2009). According to this hypothesis, the more self-related a stimulus is, the more strongly the relevant brain networks will be activated. There is considerable evidence that this is the case. For example, the default mode brain regions commonly associated with self-processing are involved in representing the traits and mental states of both the self and other people (Amodio & Frith, 2006; Gillihan & Farah, 2005; Uddin, Iacoboni, Lange, & Keenan, 2007).

Legrand and Ruby (2009) take this argument a step further, claiming that the activation of the DMN in tasks involving the self or others is a red herring, obscuring its true functions: inferential processing using information retrieved from memory. Together, Legrand and Ruby dub these functions "evaluation". Across cognitive domains, the DMN is recruited when the information currently available is insufficient to perform the task at hand. That is, additional information must be selected and reconstructed from memory and used to draw conclusions: to infer enduring traits from remembered events, to discern mental states from social cues, and so on. These evaluative processes are engaged regardless of the target of the task, whether it is the self, another, or even a non-human object. Therefore, any apparent preference for the self is incidental, due to the demands posed by the task or to one of the confounding factors listed above. Legrand and Ruby (2009) suggest that the only type of self that is uniquely realized in the brain is akin to the minimal self (though they do not use this term). However, even if there is no neural specificity for the extended self, it may still be a useful construct. Extended self-processing may be a particularly elite form of more mundane neural and mental processes.

Although my previous data suggest that the DMN and the core self network are recruited by self-relevant processing in a graded manner, I recommend caution in drawing conclusions for the following reasons. First, the values of the self-relatedness variable I used are somewhat arbitrary. Although the self is clearly more self-related than Oprah, and Oprah is more self-related than the Blue Flame, there is no reason to think that these differences are equal in magnitude. Second, I did not account for individual differences. Self-relatedness is a subjective phenomenon, not an intrinsic property of individuals. To address these concerns in the current experiment, I allowed participants to choose individuals form their own lives, and to give them subjective ratings for closeness to self. I chose closeness, as opposed to familiarity, because closeness has been shown to modulate neural activity in self-processing paradigms (Moran et al., 2010; Krienen et al., 2010).

I operationalized closeness to self with the Inclusion of Other in the Self Scale (IOS; Figure 13), a seven-point likert scale in which participants rate the closeness of some individual to oneself by choosing from seven Venn diagrams indicating various degrees of overlap (Aron, Aron, & Smollan, 1992). Besides its behavioral validation, the IOS is correlated with neural overlap between representations for self and familiar others (Beckes, Coan, & Hasselmo, 2013). I imposed a value of 8 for total overlap with oneself, an option not given in the scale. In this way, I hoped to achieve a more sensitive scaling of the construct: rather than three rather arbitrary values, I will have seven possible values that are sensitive to individual experience.

Issue #2: Do core representations for groups exist? If so, are they instantiated by the same network as core representations of individuals? My previous results (Figure 3) suggest that the brain forms core representations of individuals, and that it does so using the same neural system regardless of whether the individual is the self, another human, or a non-human object. First, I will provide a plausible framework for interpreting my initial findings. Second, I will consider the possibility that the brain also forms core representations for groups of people.

Sensory input is continuous and undifferentiated, but people perceive an environment rich with discrete events and individuated objects. From myriad wavelengths of light striking their retinas, people discern people and things; from vibrations in the air people parse words and sentences, or extract melodies from background noise. Furthermore, people perceive objects as enjoying a continuity of identity as experience plays out, even though the direct sensory evidence would lead them to believe otherwise: color, luminance, and location can vary, and objects can even disappear from sight, yet something in the way people represent objects remains invariant with respect to these changes (Scholl, 2007). Piaget's pioneering experiments revealed that humans are capable of representing object permanence as early as the second year of life, but more recent experiments have found evidence for such abilities in infants as young as 2.5 months old (Carey & Xu, 2001).

In the adult visual attention literature, object-oriented cognition is called *mid-level* because it resides in between low-level sensory processing and high-level categorization and conceptualization. This most basic representation of an object, which I argue is equivalent to what I call core representations, is known as an *object file*. Object files allow visual experience to take place in terms of people and things, rather than a patchwork of colors and contours (Gao & Scholl, 2010). Object files may be updated to incorporate the properties of that object, serving as pointers to those properties, but they can also exist independent of such conceptual information (Carey & Xu, 2001; Noles, Scholl, & Mitroff, 2005).

Object files empower cognition by freeing it from the dimensions that constrain perception. For example, Noles and colleagues (2005) demonstrated *object-specific preview benefits* (OSPBs) in a visual attention and memory task. Boxes with capital letters written inside them were presented on a screen. The letters then disappeared, but the boxes remained. Sometimes the boxes moved about the screen, but always in a manner consistent with real physical objects. Finally, a letter appeared in one of the boxes, and participants were asked to make an old/new judgment. Accurate "old" responses were faster when the letter appeared in the correct box, irrespective of position on the screen. The authors inferred that participants were forming object files that bound the letters to the boxes that contained them, creating a single mnemonically stable entity.

Therefore, at its most basic, the mental representation of an object is both pre-conceptual and robust to changes in perceptual features. I consider object files a suitable tool for understanding core representations, even though my paradigm used internally generated representations of people and objects as opposed to visually attended objects. Although the notion of object files originated in the visual attention literature, it is applicable in other domains. For example, object files are thought to underlie numerical cognition for small quantities, explaining the dissociation between subitizing and those mental computations that follow Weber's Law (Johnson & Munakata, 2005). Auditory percepts can also be processed as objects (Scholl, 2001; 2007; Zmigrod & Hommel, 2009), and there is evidence that object files can retain their identities across sensory modalities (Jordan, Clark, & Mitroff, 2010). Shen and Mondor (2008) used auditory object files in an elegant explanation of the attentional blink effect, attributing the lapse in attention to the cognitive demands of creating a new object file. Planned movements are thought to be organized in "action files", and "event files" are thought to underlie episodic memories (Hommel, 2004, Zmigrod & Hommel, 2009). The concept of object files has even been extended into the domain of social cognition and person understanding (Robbins, 2002). Indeed, object files may exist outside of any perceptual modality, representing relatively abstract information, such as identity (Gordon & Irwin, 1996).

The core self network consists of regions likely necessary for the instantiation and maintenance of object files. The large bilateral occipital clusters in the core self network are located in the ventral visual stream, beyond the striate regions associated with low-level perceptual processing, but prior to the association areas associate with complex object recognition (Nassi & Callaway, 2009). In other words, I found core representations of individuals exactly where one would expect mid-level visual cognition to be taking place. Also, core representations must be held in mind – for 3 seconds in my paradigm – before they can be used to create conceptual representations via the predication process. This is equivalent to an object file being created and maintained in working memory (Woodman & Luck, 2004), a process that would engage the ventrolateral prefrontal cortex (Postle, 2006).

In sum, I have a working framework for understanding how the brain instantiates core representations for individuals. In the current experiment, I addressed the question of whether the brain also instantiates core representations for groups, and whether these resemble core representations for individuals.

It is likely that understanding groups requires the benefits afforded by core representations. In light of the object file interpretation, core representations function as cognitive entities that are robust enough to participate in cognitive tasks, yet flexible with respect to the actual conceptual content they store. There is considerable evidence to suggest that categories are represented flexibly.

Psychological categories are not clearly defined. Even for natural categories such as *birds*, *clothing*, *vegetables*, *clothing*, and *furniture*, people exhibit typicality effects in judging category membership. For example, reaction times in category verification tasks are faster for typical category members (e.g., "is a *robin* a *bird*?") than for atypical members (e.g., "is an *ostrich* a *bird?*"). The same is true in reverse: typical category members lead to faster category naming (Sandberg, Sebastian, & Kiran, 2012). As a result, it is often difficult to define category boundaries. In one seminal study, it was found that perceived category boundaries exhibited considerable discrepancy between subjects and inconsistency within subjects from one occasion to the next (McCloskey & Glucksberg, 1978).

Some categories are created dynamically in response to particular goals (e.g., *things to take with you from a burning house*; Barsalou, 1983; 1985). Best known as "ad-hoc categories", these exhibit typicality effects much like natural categories. Whereas typicality gradients of natural categories tend to reflect the central tendency of their exemplars (e.g., *dogs* tend to be between one and three feet long and earthy in color), typicality gradients of ad-hoc categories tend to reflect those qualities that ideally realize the category's goals (e.g., *baby* is an ideal *thing to take with you from a burning house*, particularly if your goal is to save items of emotional significance). Typicality gradients for both natural kinds and ad-hoc categories reflect experiential factors such as the frequency with which particular exemplars have been encountered (Barsalou, 1985).

In short, categories are not well-behaved, discrete entities. Therefore, it is likely that groups of people must be represented dynamically and flexibly in order to facilitate their use in cognitive tasks. Core representations, as understood in the object file framework, provide a plausible mechanism.

Assuming that core representations of groups exist, do they resemble core representations of individuals? One way to frame the question is to ask whether the difference between individuals and groups is fundamental or merely syntactic, that is, a vagary of language. There is some evidence to suggest the former: infants as young as 22 months can distinguish between singular and plural sets (Li, Ogura, Barner, Yang, & Carey, 2009). In English-learning children, this ability correlates with the acquisition of linguistic distinctions between singular and plural nouns. However, children learning languages with impoverished singular-plural marking are not at a disadvantage in learning to distinguish singular and plural sets. It therefore seems likely that the brain represents plurality in some minimal fashion that does not require language. Therefore, one may expect that core representations for groups exist, and that they differ from core representations for individuals.

Issue #3: Does activity in the classic self network reflect the predication of properties upon people, or the representation of those properties per se? The classic self network is activated when people process conceptual information related to people (Gillihan & Farah, 2005). One possibility is that conceptual information about people is being explicitly represented by it. Another possibility is that conceptual information is explicitly represented elsewhere in the brain, and that the self network plays a role in binding it with core representations of the relevant individuals. My previous results are consistent with either interpretation, or a blend of the two. One goal of the present experiment was to disentangle competing explanations on a region-by-region basis.

The medial prefrontal cortex and posterior cingulate cortex are consistently activated by paradigms in which traits, faces, actions, and other stimuli are processed with respect to the self (Mitchell, 2009; Powell et al., 2010). These activations are revealed by contrasting self-processing conditions against conditions in which stimuli are processed with respect to other people or in some surface-level manner. Cortical midline structures involved in self-representation also show a preference for autobiographical memory retrieval tasks, presumably by representing the self as situated in a remembered context (Northoff & Bermpohl, 2004; Svoboda, McKinnon, & Levine, 2006).

Accordingly, patients with frontal lobe damage can have severe alterations in self, even when other functions remain normal. One neurological review tells of a “charming, socially dynamic” patient who became “irritable, aggressive, and domineering”. Other cases in the same review are similar: frontal lobe degradation results in drastic changes in personality and self-schemas (Miller et al., 2001).

The cortical midline, particularly the medial prefrontal cortex, is consistently involved in and even necessary for the conceptual self. Still, it is not clear that these regions are sufficient for extended self-representation, or even that they perform a representation function per se. One possibility is that properties are represented elsewhere, and classic self regions bind them to a minimal self-representation. More generally, the same principles could apply to representations of other people. In other words, the cortical midline and perhaps the TPJ would be engaged in a predication process that unifies minimal and conceptual representations. Such a process would be distinct from the explicit representation of either (Carey & Xu, 2001).

Classic self regions are activated by a diverse set of cognitive tasks that is broader than, but likely involved in, the representation of individuals. For example, the areas in the DMN are involved in memory recall, inference, and evaluative judgments (Legrand and Ruby, 2009). All of these processes are critical in making judgments about an individual’s traits, allowing for the possibility that those traits are explicitly represented elsewhere in the brain, while the cortical midline is involved in linking those conceptual representations with core representations of the appropriate individual. This idea has interesting parallels with the somatic marker hypothesis for affective decision making. According to Damasio (1996), the ventromedial prefrontal cortex facilitates emotional processing by binding semantic and perceptual information with bodily states. It is a small conceptual leap to suppose that the ventromedial prefrontal cortex likewise binds embodied core representations of individuals with semantic information to construct conceptual representations.

If properties are represented explicitly in the classic self network, then they should elicit cortical midline and TPJ activation even when they are not being predicated of an individual. On the other hand, if these activations reflect a predication process by which properties are bound to individuals, then properties should elicit the associated activity only when they are being predicated of an individual. To test these hypotheses, I employed a task that requires the mental representation of a property without an individual upon which to predicate it. Specifically, participants made semantic judgments about the properties themselves. In some trials, participants decided whether a property is mental or physical. In other trials, they decided whether a property is learned or genetic.

Overview

The current experiment intended to address the three issues outlined above in addition to replicating my previous experiment. To that end, each functional run was divided into blocks, each consisting of a somewhat different task designed to address a particular issue (Figure 10). Blocks 1 and 2 addressed issue #1, block 3 addressed issue #2, and block 4 addressed issue #3. It should be noted that the actual block order was counterbalanced between runs.

Although the three tasks were somewhat different, the structure of a trial remained identical to the one used in my previous experiment (Figure 11): two stimuli were presented consecutively for 3 seconds each. During the presentation of the second stimulus, a response was required. One-quarter of the trials were *catch trials*, meaning that the second stimulus was omitted and no response was required. These catch trials allowed for the linear separation of the hemodynamic response to the first stimulus from the hemodynamic response to the second (Ollinger et al., 2001a; 2001b).

Blocks 1 and 2: the individuals task. In blocks 1 and 2 participants were presented with an *individual* and a *property* (Figure 11), displayed as one or more printed words on a computer screen, and were asked to rate how well the property applies to the individual. The four individuals used were the participant herself, a relative of hers, a friend, and an acquaintance. The three possible responses were "applies poorly", "applies somewhat", and "applies well". Blocks 1 and 2 each only featured two of the four individuals, in order to maintain consistency with the other two tasks.

Blocks 1 & 2 were essentially a replication of my previous experiment using different individuals, so I predicted similar results. First, I predicted that the conjunction of all four *individuals* conditions would reveal the identity network (Figure 3). Second, I predicted that a linear amplitude-modulated regression in the *individuals* conditions, using closeness as measured by the IOS as the parameter, would reveal a set of regions adjacent to the core self network (Figure 7), as well as the extrastriate body area bilaterally (Figure 8). Third, because all four individuals in the current experiment are humans, the conjunction of all four *properties* conditions from blocks 1 & 2 would reveal the ventromedial and dorsomedial prefrontal cortex, the posterior cingulate, and the TPJ (Figure 5). Finally, a linear amplitude-modulated regression in the *properties* conditions, using IOS scores as the predictor, would reveal the classic self network (Figure 9). The two linear regressions serve to address issue #1 by employing a more sensitive measure of self-relatedness than my previous experiment.

Block 3: the groups task. In block 3, participants were presented with a *group* and a *property* (Figure 11), displayed as one or more printed words on a computer screen, and were asked to rate how well the property applies to the group. The two groups used were “adults” and “children”. The three possible responses were “applies poorly”, “applies somewhat”, and “applies well”. Block 3 was intended to address issue #2. If core and conceptual representations for groups exist, then the activations associated with *group* conditions should reveal the network used to instantiate them.

Block 4: the semantic task. In block 4, participants were presented with a semantic *prompt* and a *property* (Figure 11), displayed as one or more printed words on a computer screen. The two prompts used were “Genetic” and “Physical”. Participants were asked to answer the following question: “To what extent is the following property genetic/physical?”. Block 4 was intended to address issue #3. If activity in the DMN reflects the predication of properties upon individuals, then activation should be significantly lower for properties in block 4 than in the other blocks. If properties are explicitly represented in the DMN, then they will activate the DMN even when they are not being predicated of an individual.

Methods

Participants

Nineteen right-handed, native English speakers with normal hearing and normal or corrected vision were recruited from the Emory community. One additional participant was excluded due to extreme head motion in the scanner. Participants received $80 and images from their anatomical scans.

Design

A repeated-measures design included three variables: (1) Phase, (2) Subject, and (3) Subject Property. For Phase, participants received an individual, group, or prompt during Phase 1, and then received a property during Phase 2. For Subject, participants received one of the eight following subjects on each trial during Phase 1: the name for one of four individuals ("I", "Relative", "Friend", or "Acquaintance," the word for one of two groups ("Adults" or "Children), or one of two semantic prompts ("Physical", or "Genetic"). I refer to these eight stimuli as "subjects" because each is the subject, syntactically speaking, of the predication and judgment that followed. As can be seen, the eight subjects were nested within Individuals (4), Groups (2), and Prompts (2) (i.e., not crossed). For Subject Property, participants evaluated exactly the same set of 36 properties for each of the eight subjects during Phase 2. Thus, Subject Property had eight levels, one for each of the eight subjects that occurred during Phase 1, again with the properties held constant (i.e., the value of Subject during Phase 1 specified the value of

Subject Property during Phase 2). After receiving the property, participants rated how well the property applied to the subject that preceded it. Thus,the design took the form of a fully within-subjects factorial design, with eight subjects crossed with two phases. The experiment was divided into six functional runs, each containing four blocks. Within a block, only one of the three tasks was featured (as described later). In this way, every condition was equally represented within each run.

Materials

All stimuli in the individuals task, groups task, and semantic task consisted of words or short phrases presented in the center of the screen. For subjects, the materials included the words I, Relative, Friend, Acquaintance, Adults, Children, Physical, and Genetic. For the properties, the materials contained the 36 properties shown in Table 12. Fixation was a white cross on a black background with no border (Figure 11).

Individuals, groups, and semantic prompts. Individuals, groups, and semantic prompts were presented to participants as single words: white text against a black background in 18-point Georgia font. To ensure that gaze remained consistent, with attention directed towards a constant angle of view, all stimuli were surrounded by a white border of fixed size. Prior to scanning, participants were instructed in detail as to what each of the 8 subjects represents (“I”, Relative”, etc.), so that during scanning, they were able to process the correct individual, group, or prompt deeply, and to predicate the corresponding property upon it.

Properties. Much like the subjects, properties were presented to participants as typed words in white 18-point Georgia font on a black screen, surrounded by the same white border. The same set of 36 properties was used for all individuals and groups and in the semantic task (Figure 12).

Half of the properties used were physical (i.e., concrete, perceptible to the senses), and the other half were mental (i.e., abstract, psychological, social, personality related), a distinction I will refer to as “class.” Each class was subdivided into 3 “sub-classes” as described next. Properties were counterbalanced across these sub-classes for word length, number of words, and part of speech.

The sub-classes assigned to properties differed depending on whether they were physical or mental (Figure 11). Physical properties were designated as *visuospatial*, *auditory*, or *sensorimotor*, in order to provide a similar level of diversity as in the mental properties. Mental properties were designated as being related to *conscientiousness*, *agreeableness*, or *openness to experience*: three dimensions of the Big Five personality space (McCrae & Costa, 2003).

Sub-classes for physical properties were relatively straightforward, pertaining largely to the sensory modality used in judgments regarding those properties. Visuospatial properties were those that can be evaluated according to simple visual or spatial criteria (e.g., “tall” and “long arms”). Auditory properties were those that can be evaluated according to simple auditory criteria (e.g., “soft voice” and “giggling”). Sensorimotor properties were those that relate to spatial structure or motion (e.g., “good posture” and “nimble”).

Sub-classes for mental properties were more subtle. I derived them from McCrae and Costa (2003), a study of human personality and its variable characteristics. *Conscientiousness* refers to such traits as ethical behavior, responsibility, and dependability (e.g., “self-discipline” and “rational”). *Agreeableness* refers to such traits as sympathy, compassion, amicability, and generosity (e.g., “friendly” and “eager to please”). *Openness to experience* refers to such traits as nonconformity, adventurousness, and rebelliousness (e.g., “vivid imagination” and “spontaneity”).

Note that the sub-classes as described above were used solely as a counterbalancing measure; this was not an experimental manipulation because there were insufficient trials to achieve sufficient power in BOLD signal contrasts.

Beginning with 16 properties from each subclass, I collected ratings for how “physical” (vs. mental) and “genetic” (vs. learned) each property was. I then carefully chose a subset of 36 properties representing a uniform distribution of physicalness and geneticness. All of the original subclasses were represented, but not equally. See Appendix 1 for details regarding the norming study.

Tone. The tone used in the baseline task was a pleasant synthetic orchestral string sound. It was a C-Major chord comprised of C2, E2, G2, and C3 that persisted for 500 milliseconds with natural-sounding attack and release.

Procedure

During our correspondence prior to the experiment, participants were asked to provide three individuals: a relative, a friend, and an acquaintance. In the training session, participants provided informed consent and were screened for a battery of potential problems that could arise during a magnetic resonance scan. They were then presented with their three chosen individuals and asked to rate them on the Inclusion of Other in the Self Scale (Figure 13), a validated 1-7 scale of closeness to self (Aron, Aron, & Smollan, 1992). The anticipated order (relative > friend > acquaintance) was preserved in 10 of the 19 participants. Because my imposition of the self-relatedness scale in the previous study was a possible limitation, I decided to allow participants' judgments to stand in the later analyses, even when they violated my expectations. If, for example, participants perceived their friend as closer than their relative, self-relatedness was assumed to be higher for the friend.

Participants received detailed instructions for the individuals task, groups task, semantic task, and baseline task on a personal computer running an interactive E-Prime script. During these instructions and the subsequent practice run, participants used an E-Prime button box to practice making responses. In the scanner, participants used a Current Designs fiber optic button box designed for high magnetic field environments. Using the same individuals, groups, and semantic prompts, and properties similar to but different from those actually used in the experiment, participants were gradually acclimated to the paradigm. Specifically, participants successively received practice on the individuals task, the individuals task with catch trials, the groups task, the semantic task, the baseline task, and finally all tasks and trial types together. The parallel structure of the three main tasks was emphasized to facilitate participants' acclimation to the paradigm. The experimenter was present in the room to answer questions, which were rarely though sometimes asked. Besides imparting the details of the paradigm, instructions emphasized the importance of thinking deeply about the individual/group/prompt as soon as it was presented, as opposed to waiting for the property.

Next, at the same computer, participants performed a full practice run of 48 complete trials and 12 catch trials. Again, the properties received were similar to but different from those that were later used in the actual experiment. The experimenter was not present in the room. When the participant emerged about 5 minutes later, he/she was asked to confirm that he/she understood the task and was ready to perform it in the scanner.

Following the practice run, the experimenter and the participant walked downstairs to the FERN imaging center. Once settled safely and comfortably in the scanner, an initial anatomical scan was performed, followed by the six critical functional runs. Two brief localizer runs were then performed to identify cortical regions specifically recruited in the visual processing of faces, bodies, and words. The localizer consisted of a 1-back task in which participants pressed the button when they saw the same image twice. Blocks of the task used images of faces, bodies, pseudowords, or objects (as a control for low-level visual processing). I am grateful to Dr. Danny Dilks for providing the stimuli and advice on experimental design for the localizers.

Total time in the scanner was approximately 1 hour and 15 minutes. Over the course of the scanning session, participants received 384 trials. On 25% of these trials, the presentation of the property was omitted. These are called catch trials, and their inclusion in the experimental design permits linear separation of the BOLD responses to individuals/groups/prompts from those to subsequent properties. In a catch trial, the participant was not asked to make a response; the presentation of the individual/group/prompt was immediately followed by a return to the baseline task. Trials that included both an individual/group/prompt and a property are called complete trials, whereas those including only an individual/group/prompt are called catch trials. Of the 384 total trials, 288 were complete trials and 96 were catch trials. The catch trials constituted 25% of the total trials, a proportion in the recommended range for an effective catch trial design (Ollinger et al., 2001a; 2001b).

The trials were distributed across 6 runs. Each run lasted about 9 minutes and 30 seconds and contained 36 complete trials and 12 catch trials. Both types of trials (complete and catch) were randomly intermixed as described earlier. On a given trial, participants were not able to predict whether a complete trial or a catch trial was coming, a necessary condition for an effective catch trial design (Ollinger et al., 2001a; 2001b). Random ISI ranging from 3 to 12 sec (in increments of 3 sec) occurred between trials. Each individual/group/prompt occurred 8 times per run, followed 6 times by a property (complete trial) and twice by a return to fixation (catch trial).

Four blocks occurred per run in carefully counterbalanced order: two blocks of the individuals task, one block of the groups task, and one block of the semantic task (Figure 10). Over the course of 6 runs, every possible order was used that did not feature consecutive individuals task blocks. In each block of the individuals task, two of the individuals (“I”, “Relative”, “Friend”, or “Acquaintance”) were presented 8 times per block for 3 seconds, for a total of 48 presentations of each individual across the 6 runs. In each block of the groups task, each of the two groups (“Adults” and “Children”) was presented 8 times per block for 3 seconds, for a total of 48 presentations of each group across the 6 runs. In each block of the semantic task, each of the two prompts (“Physical” and “Genetic”) was presented 8 times per block for 3 seconds, for a total of 48 presentations of each prompt across the 6 runs.

Each of the 36 properties appeared twice per run: once in the semantic task, and once in one of the other two tasks. These assignments were meticulously counterbalanced and pseudorandomized. Therefore, on a given trial, participants were unable to predict either the subject or the property class. In each block, the critical trials and catch trials were ordered in a rapid event-related design according to one of 48 sequences selected out of 100,000 as optimal by the optseq2 algorithm (http://surfer.nmr.mgh.harvard.edu/optseq).

Any complete trial lasted for 6 seconds, not including a variable intertrial interval. The individuals, groups, and semantic tasks were identical in structure (Figure 10). First, an individual (“I”, “Relative”, “Friend”, or “Acquaintance”), group (“Adults” or “Children”), or prompt (“Physical” or “Genetic”) was presented for 3 seconds. Participants were instructed to think about the individual/group/prompt in order to prepare for the presentation of the property. Second, one of 36 properties was presented for 3 seconds. During the property presentation, participants were instructed to answer the question “How well does the property apply to the individual/group?” or “To what extent is the following property physical/genetic?” using one of three buttons on the fMRI button box.

In between trials of the three main tasks, participants fixated on a small white cross in the center of the black screen. For the duration of this intertrial interval, participants were instructed to remain vigilant for a tone, and to respond with a button press upon hearing it. The tone occured once during every intertrial interval, but the onset time varied randomly from 500 ms after the end of the previous trial to 1500 ms before the start of the subsequent trial. This procedure constitutes the baseline task.

Image acquisition and analysis

The neuroimaging data were collected at the Facility for Education and Research in Neuroscience (FERN) at Emory University on a research-dedicated 3T Siemens Trio scanner. In each functional run, T2*-weighted echo planar image volumes depicting BOLD contrast were collected using a Siemens 12-channel head coil and parallel imaging with an iPAT acceleration factor of 2. Each volume was collected using a scan sequence with the following parameters: 56 contiguous 2 mm slices in the axial plane, interleaved slice acquisition, TR = 3000 ms, TE = 30 ms, flip angle = 90°, bandwidth = 2442 Hz/Px, FOV = 220 mm, matrix = 64, voxel size = 3.4 mm × 3.4 mm × 2.0 mm. This scanning sequence was selected after testing a variety of sequences for susceptibility artifacts in orbitofrontal cortex, amygdala, and the temporal poles. I selected this sequence not only because it minimized susceptibility artifacts by using thin slices and parallel imaging, but also because using 3.4 mm in the X-Y dimensions yields a voxel volume large enough to produce a satisfactory temporal signal-to-noise ratio.

In the anatomical run, 176 T1-weighted volumes were collected using a high resolution MPRAGE scan sequence with the following parameters: 192 contiguous slices in the sagittal plane, single-shot acquisition, TR = 2300 ms, TE = 3.02 ms, flip angle = 8°, FOV = 256 mm, matrix = 256, bandwidth = 130 Hz/Px, voxel size = 1.0 mm × 1.0 mm × 1.0 mm.

Image preprocessing

Image preprocessing and statistical analysis were conducted in AFNI (Cox, 1996). Slice-time correction was performed on the functional volumes followed by motion correction and registration to the participant’s anatomic image, which were performed in a single step to reduce rounding error that occurs when the functional data are independently warped multiple times. A transformation matrix for motion correction was generated in which all functional volumes were registered to the first volume (in which anatomical information is thought to be best preserved) of the first run (which is nearest in time to the acquisition of the anatomical image).

To generate the transformation matrix form native space into normalized Talairach space, the anatomical image was skull-stripped and aligned to the same functional volume used as the registration base for motion correction. Voxels outside the brain were removed from further analysis, as were high-variability low-intensity voxels likely to be shifting in and out of the brain as a result of minor head motion. The skull-stripped and aligned anatomical image was then transformed to Talairach space using an automated procedure employing the TT_N27 template (also known as the Colin brain, an averaged dataset from one person scanned 27 times). The matrices generated from the Talairach transformation of the anatomical dataset were set aside at this point.

The functional data were then smoothed using an isotropic 8 mm full-width-half-maximum Gaussian kernel. Finally, the signal intensities in each volume were divided by the mean signal value for the respective run and multiplied by 100 to produce percent signal change from the run mean. All later analyses were performed on the percent signal change data. In all later regressions, the six regressors obtained from motion correction during preprocessing were included to remove any residual signal changes correlated with movement (translation in the X, Y, and Z planes; rotation around the X, Y, and Z axes). Scanner drift was modeled by finding the best-fitting polynomial function correlated with time in the preprocessed time course data for each run. After individual-level analysis using the general linear model described below, the voxel dimensions of the functional volumes were resampled from 3.4 mm × 3.4 mm × 2.0 mm to 3.0 × 3.0 × 3.0 mm, and warped into Talairach space. Preprocessing was identical, but performed separately, for the localizer runs.

Data analysis: individual level

Regression analysis was performed at the individual level using a canonical, fixed-shape gamma function to model the hemodynamic response. The gamma function was convolved with stimulus onset times to produce a regressor for each of 16 conditions (plus one regressor of non-interest for the instructions presented at the beginning of each block). Specifically, betas were calculated each of the following conditions, for each voxel in the brain: four individuals (“I’, “Relative”, “Friend”, and “Acquaintance”), two groups (“Adults” and “Children”), 2 prompts (“Physical” and “Genetic”) and the eight property conditions associated with each. The individuals/groups/prompts conditions will from now on be referred to as “Phase 1”, and the associated property conditions will be referred to as “Phase 2.” A subscript indicating Phase 1 or Phase 2 will be used when referring to each condition.

For later analyses, I combined all of the individuals into a single regressor for each phase. For example, I_1, $\textit{Relative}_1$, $\textit{Friend}_1$, and $\textit{Acquaintance}_1$ were summed into $\textit{Individuals}_1$. The same procedure was repeated for Phase 2, and for groups and prompts, yielding six additional regressors ($\textit{Individuals}_1$, $\textit{Groups}_1$, $\textit{Prompts}_1$, $\textit{Individuals}_2$, $\textit{Groups}_2$, and $\textit{Prompts}_2$).

I also performed analyses in which I was interested in the effect of closeness to self (as measure by the IOS) within the four individuals conditions during each phase. Here, I created amplitude-modulated regressors in which the gamma function, prior to convolution with stimulus onset times, was multiplied by the appropriate IOS score. The regression beta coefficient for such a regressor may be treated like Pearson's *r*, where positive/negative values indicate a positive/negative association between hemodynamic response and closeness to self.

For the localizers, which followed a blocked experimental design, regressors were generated using AFNI's "block" function for each 15-second block of the 1-back task for faces, bodies, words, and objects. Intermittent periods of fixation constituted an unmodeled implicit resting baseline.

As described earlier, the catch trial design allowed us to separate hemodynamic responses for the individuals/groups/prompts from those for the subsequent properties that followed immediately without random jitter. The Phase 1 conditions were modeled by creating regressors including presentations from both complete trials and catch trials. Including individuals, groups, and prompts from both trial types in one regressor make it possible to linearly separate them from the subsequent property conditions. Thus, activations from Phase 1 conditions were not conflated with the activations for the Phase 2 properties conditions.

Data analysis: group level

The beta coefficients for the 16 conditions from each participant's regression were entered into group analyses using repeated-measures ANOVA, in which participants were treated as random effects, and condition betas were treated as fixed effects. A spatial extent threshold was established using 3dClustSim in AFNI to estimate extent thresholds needed to correct for false positives resulting from multiple comparisons. Based on residual smoothness (dependence between voxels) and a whole-brain mask, 3dClustSim uses a Monte Carlo simulation to determine the number of contiguous voxels required per cluster to satisfy user-defined significance thresholds at the voxel and whole-brain level. I chose a voxel-wise significance level of $p < .005$ and a brain-wise significance level of $p < .05$ based on arbitrary standards established in the literature.

Statistical analyses fall into three broad types: *condition* analyses reflect individual conditions (including amplitude-modulated regressions) compared against the active baseline. *Contrasts* reflect the differences between conditions, and *conjunctions* reflect the overlap between conditions. I perform each type of analysis for Phases 1 and 2, and at two different levels: the collapsed level of *Individuals*, *Groups,* and *Prompts*, and the specific level of *I*, *Relative*, *Friend*, and *Acquaintance*.

Within each level (specific persons and collapsed by task) and phase (Phase 1 and 2), I performed every possible contrast, for the sake of completeness. For the conjunctions, I report every possible conjunction of conditions or their negations. For example, "*Individuals*$_1$ and *Groups*$_1$ and *not Prompts*$_1$", reported as "*Individuals*$_1$ * *Groups*$_1$" for simplicity, indicates clusters in which *Individuals*$_1$ and *Groups*$_1$ are significant, but *Prompts*$_1$ is not. Because I did not calculate inferential statistics for the conjunctions (only for the conditions going into them), I chose an arbitrary extent threshold of 20 voxels for these tables. The AM-regression on closeness using the IOS was treated like any other condition: betas from the individual level analysis were entered into a t-test where each voxel's beta value was compared against zero. Together, the condition, contrast, and conjunction analyses address the issues raised above.

Results

After briefly describing results related to the replication of my previous experiment (Drucker et al., submitted), the remainder of the results section is organized according to the scheme in Figure 14. First, results from Phase 1 are presented that assess issues related to core representations at the individual person level. I present the 4-way, 3-way, and 2-way conjunctions between I_1, $\mathit{Relative}_1$, Friend_1, and $\mathit{Acquaintance}_1$ (Table 1), followed by pairwise contrasts between each of those conditions (Table 2). Table 3 shows the brain areas whose activation correlates either positively or negatively with closeness as measured by the IOS. Core and conceptual representations (Phase 1 and Phase 2, respectively) are presented separately.

Next, I present the results of analyses in which conditions were collapsed by task. At this collapsed task level, the three conditions in each phase are *Individuals* (I, Relative, Friend, and Acquaintance), *Groups* (Children and Adults), and *Prompts* (Genetic and Physical). Just like at the individual person level, I present the 3-way and 2-way conjunctions (Table 4) followed by the pairwise contrasts between conditions (Table 5).

Phase 2 results are then presented that assess issues related to conceptual representations. Phase 2 results are organized in the same way as Phase 1 results. At the individual person level, I present conjunctions between I_2, $\mathit{Relative}_2$, Friend_2, and $\mathit{Acquaintance}_2$ (Table 6), followed by pairwise contrasts between each of those conditions (Table 7). Table 8 and Table 9 respectively present the conjunctions and contrasts at the collapsed task level.

Finally, I contrasted the self (I) with the three non-self other person conditions in each phase (Relative, Friend, and Acquaintance). These self versus other contrasts are presented in Table 10 and Table 11 for core (Phase 1) and conceptual (Phase 2) representation, respectively.

In some of the analyses, significant activations occurred in white matter regions of the brain. Task-related white matter activation has been demonstrated using fMRI (Mazerolle et al., 2010), but because interpretation of such activations is complicated and remains controversial, I do not address them here. Nevertheless, I include the clusters in Supplemental Table 1 for interested readers, and invite you to consider an excellent discussion of the physics and physiology underlying these issues (Gawryluk, Mazerolle, & D'Arcy, 2014).

Replication of Previous Findings

Although a different set of non-self individuals was used than in my previous experiment (Relative, Friend, and Acquaintance instead of Oprah and Flame), I predicted that the core self network identified in my previous study would be active during the Phase 1 individuals conditions of the current experiment. My previous findings, however, were only partially replicated. Figure 15 shows that for the five frontal regions in the core self network, the Phase 1 conditions were at or below baseline. However, visual occipital areas, corresponding with BA 17, 18, and 19 bilaterally, together with the left fusiform gyrus, were activated in all conditions. This foreshadows the major role that ventral visual areas play in generating core representations generally in later analyses. As Figure 15 further shows, all seven areas were activated above baseline during Phase 2, replicating the results of my previous experiment. In addition, activations were higher during Phase 2 than during Phase 1, providing further replication.

Figure 16 presents the activations for each condition in the current experiment within the cortical midline areas that became active in the previous study when generating conceptual representations (Figure 6). For non-self individuals, all four areas were activated. Strangely, vmPFC was not significantly activated by I_2, although it trended towards significance for $Adults_2$ and $Children_2$. Interestingly, retrosplenial cortex was particularly active during Phase 2 for the non-self conditions. As will be seen, retrosplenial cortex played a key role in generating conceptual representations for non-self persons (Figure 24).

Core Representations

By analyzing the Phase 1 conditions (the subjects upon whom properties were predicated), I examined core representation from several different perspectives, and at several levels. I begin with core representations for the self and other persons, working my way outward to persons more generally and groups of people.

Core representations for the self and other persons. At the individual person level, I operationalized core representations in three complimentary ways, each method providing advantages and disadvantages relative to the others. First, a conjunction analysis reveals brain activity specific to the self, other persons, and various combinations thereof (Table 1). Second, contrast analyses reveal brain activity preferential for every individual compared with every other individual (Table 2). Third, amplitude-modulated regressions reveal brain activity correlated with the closeness of an individual to oneself (Table 3).

Conjunctions. *I* first established activations for the I_1, $Relative_1$, $Friend_1$, and $Aqcuaintance_1$ conditions that were significantly higher than the active baseline, and then entered all four into a conjunction analysis. This analysis identified clusters significantly more active than baseline in every possible combination of the four conditions. Together, this set of activation maps provided us with one perspective on the neural correlates of core representation. Because Phase 1 of the individuals task ("*think deeply about the individual presented*") was designed to be as different as possible from the active, implicit baseline task ("*press the button when you hear the tone*"), one can reasonably infer that most of the brain areas involved in core representation for the four individuals will be present. Importantly, however, brain areas *not* associated with core representation but associated with the task could have potentially become active for a condition (e.g., brain areas associated with visual processing of the stimulus, or preparing strategically in anticipation for Phase 2, etc.). Because task-related activations should be active across all four individuals, any such activations should only occur in the four-way conjunction, and drop out in the contrasts. Thus, care must be taken in viewing activations in this particular conjunction. Although some could be related to core representation across all four conditions, others could be related to the task. For all other smaller conjunctions, any activations observed should only be related to core representations, again because all task activations should be common across conditions, and thus only emerge in the four-way conjunction.

Table 1 shows activations for every combination of individuals, beginning with more inclusive combinations, and working its way down to activations unique to each individual. Emerging themes include visual processing in posterior and ventral occipitotemporal regions, and spatial processing in lateral and medial parietal areas. At the single condition level (Table 1), somatosensory and motor processing appear to contribute as well.

Contrasts. To provide another perspective on core representation, I performed contrasts between each pair of conditions. Because the task was held constant across all conditions, save for the identity of the individual, one can reasonably infer that any significant clusters played a role in representing these individuals (i.e., not merely activated due to the particulars of the task). Notably, however, clusters central to core representations could be subtracted out in one or both conditions, thereby making the contrast a conservative assessment of core representations (i.e., when both conditions have activations above baseline that do not differ from each other, both are subtracted out; when both conditions have activations above baseline that do differ from each other, the lower one is subtracted out). Together, the conjunction and contrast analyses provide a broad view of the brain areas associated with core representations.

Table 2 presents the results of these contrasts. All possible pairs of individuals are shown, with clusters significantly more active for every individual relative to every other individual. Again, a continuing theme is that visual, spatial, and sensorimotor processing appear central to core representations. In general, more areas seem to have been preferentially activated for the self over other persons, rather than vice-versa, with an interesting exception in medial parietal cortex.

Amplitude-modulated regression. Finally, I performed an amplitude-modulated regression to identify brain regions involved in representing individual persons generally, but involved to a greater or lesser degree depending on how close the relevant individual is to the self. Each participant rated their closeness with each of the individuals (*I* was given an automatic rating of 8 on the 1-7 scale) using the Inclusion of Other in the Self Scale (IOS). Supplementary Table 2 presents the ratings for each individual from each participant. Note that typically I > Relative > Friend > Acquaintance, but I allowed exceptions (e.g., Participant #14) so that the data would be driven by participants' personal feelings, not my own hypotheses. I created an amplitude-modulated regressor for each participant, where the amplitude of the convolution kernel (the gamma function used to estimate the hemodynamic response) was multiplied by the participant's rating for the individual presented on that trial. This amplitude-modulated regressor isolated variance associated with the closeness of the individual to oneself during Phases 1 and 2 (c.f., Drucker et al., submitted, in which a similar regressor was created using coefficients chosen a priori). As in the previous experiment, positive correlations imply that activation increased with closeness to oneself, and negative correlations imply that activation decreased with closeness to oneself.

Table 3 shows clusters of positive correlation in late visual areas and in the basal ganglia, together with negative correlation in earlier visual areas. The negative cluster should be interpreted with caution, because it could be an artifact of low-level visual properties of the stimulus (see Supplementary Figure 1 – Supplementary Figure 8). Additionally, a few interesting subthreshold clusters were observed in the left dorsomedial prefrontal cortex, temporoparietal junction, and anterior temporal lobe, all areas implicated in the processing of person knowledge.

Core representations for individual persons and groups of people.

I collapsed conditions by task in order to address Issue #2: how representations for groups of people are similar or different from representations for individual persons.

I$_1$, *Relative*$_1$, *Friend*$_1$, and *Acquaintance*$_1$ were collapsed into *Individuals*$_1$; *Adults*$_1$, and *Children*$_1$ were collapsed into *Groups*$_1$; and *Physical*$_1$, and *Genetic*$_1$ were collapsed into *Prompts*$_1$. As this collapsed task level, much like at the specific person level, I operationalized core representations in two complimentary ways. First, a conjunction analysis reveals brain activity specific to individuals, groups, the semantic prompts, and the various combinations thereof (Table 4). Second, contrast analyses reveal brain activity preferential for individuals, groups, and prompts compared with one another (Table 5). Amplitude-modulated regressions were not appropriate at this level.

Conjunctions. I compared the *Individuals*$_1$, *Groups*$_1$, and *Prompts*$_1$ conditions against the baseline model, and entered all three into a conjunction analysis. The conjunction identifies clusters significant against baseline in every possible combination of the three collapsed conditions. This provided us with a liberal definition of the neural correlates of core representation at this level: because Phase 1 of the individuals and groups tasks ("*think deeply about the individual/group presented*") was designed to be as different as possible from the active, implicit baseline task ("*press the button when you hear the tone*"), one can reasonably infer that most of the brain areas involved in core representation for individuals and groups of people will be present.

Table 4 shows activations for every combination of Phase 1 conditions, with the three-way conjunction (Table 4), two-way conjunctions excluding the third condition (Table 4), and finally each condition in isolation (Table 4). Unsurprisingly, many of the visual, spatial, and sensorimotor regions from the specific-level analysis are present for the individuals. For prompts, many unique activations, including several in medial frontal regions associated with the default mode network, warrant later discussion. For groups, there appears to be little activity outside of visual areas.

Contrasts. To provide a more conservative measure of core representation, and to see how the Phase 1 individuals, groups, and prompts related to one another, I performed contrasts between each condition. Because the task was identical in structure across conditions, one can reasonably infer that any significant activations were due to differences in core representation.

Table 5 shows the results of these contrasts. All possible pairs of conditions are shown, with clusters significantly more active for every condition relative to every other condition. Compared with groups, individuals activated motor planning regions and lateral parietal cortex, with interesting implications for body ownership and agency to be discussed later. Intriguingly, there were no significant activations for either individuals or groups compared with prompts. Compared with individuals and groups, the semantic prompts activated regions involved in visual processing and executive function, as well as a few subcortical regions.

Conceptual Representations

By analyzing Phase 2 conditions (the properties), I examined conceptual representation from several different perspectives, and at several levels. I begin with conceptual representations for the self and other persons, working my way outward to persons more generally and groups of people.

Conceptual representations for the self and other persons. As with core representations earlier, I operationalized conceptual representations for the self and other persons in three complimentary ways: a conjunction analysis (Table 6), contrast analyses (Table 7), and an amplitude-modulated regression (Table 3).

Conjunctions. *I* compared the I_2, $Relative_2$, $Friend_2$, and $Acquaintance_2$ conditions against the baseline model, and entered all four into a conjunction analysis. The conjunction identifies clusters significant against baseline in every possible combination of the four conditions. Table 6 shows activations for every combination of individuals, beginning with more inclusive combinations, and working its way down to activations unique to each individual. A large left-lateralized frontal area is active in all four conditions, as are several cortical midline structures including several in the DMN. Visual areas again appear to play a prominent role.

Contrasts. To provide a more conservative measure of conceptual representation, I performed contrasts between each Phase 2 condition. Because the properties were exactly the same in every condition, one can reasonably infer that any significant clusters were involved in conceptual representation per se (i.e., not merely activated due to the particulars of the task).

Table 7 shows the results of these contrasts. A general emerges with a large set of regions preferential for the self, including left-lateralized frontal regions, bilateral parietal cortical regions, and several DMN regions in the cortical midline. Noticably preferential against the self is the retrosplenial cortex, providing a dorsal/ventral double dissociation between self and other persons in the medial parietal lobes. Other areas with more complex roles include lateral temporal cortex and visual areas.

Amplitude-modulated regression. Finally, I performed an amplitude-modulated regression in Phase 2 similar to the one performed for Phase 1. The HDR for properties were modulated according to the IOS score for the individual immediately preceding it in order to isolate variance in brain activity related specifically to closeness to self of the individual being conceptually represented. Table 3 shows positive correlation in the dorsal medial frontal lobe and the left supramarginal gyrus. Negative correlations include visual areas and medial parietal regions. The dissociation within the DMN (dmPFC vs. PCC) implies a heterogeneity in DMN function, which will be discussed later.

Conceptual representations for individual persons and groups of people.

I collapsed conditions by task in order to address Issue #2 in Phase 2, just as in Phase 1. Here, a conjunction analysis reveals brain activity specific to properties for individuals, groups, and the semantic prompts, and the various combinations thereof (Table 8). Next, contrast analyses reveal brain activity preferential for individuals, groups, and prompts compared with one another (Table 9). Amplitude-modulated regressions were not appropriate at this level.

Conjunctions. Table 8 shows the clusters in every possible combination of $Individuals_2$, $Groups_2$, and $Prompts_2$. Many of the same areas from the analogous analysis at the specific level are unsurprisingly present for properties-for-individuals. Little is unique for properties-for-groups, and the pattern implies that overlapping clusters for the three conditions tend to be more spatially extensive for $Individuals_2$ and, to a lesser extent, $Prompts_2$.

Contrasts. Table 9 shows the contrasts between Phase 2 conditions, providing considerable insight into issues raised in the introduction pertaining to how groups are represented in the brain (Issue #2) and the role of the DMN in representing individuals per se, or properties more generally (Issue #3). Numerous frontal lobe areas are preferential for *Individuals*$_2$ > *Groups*$_2$ but not vice-versa, with implications for Issue #2. Activation preferential for *Individuals*$_2$ > *Prompts*$_2$ is extensive as well, and the heavy overlap with the DMN has profound implications for Issue #3. Prompted properties preferentially activate a large set of regions as well, as can be seen in Table 9.

Localizers

During the localizer task, participants performed a 1-back task with pictures of faces, bodies, or objects (a pseudowords condition was not analyzed). The contrast *Faces* > *Objects* identified voxels selective for the visual processing of faces; the contrast *Bodies* > *Objects* identified voxels selective for the visual processing of bodies. For faces, there was a high degree of spatial variability across participants, so ROIs were identified by thresholding the resulting t-maps at ($p < .005$) at the individual participant level. In this manner, I was able to produce masks for several well-established face-selective regions in the brain. For bodies, there was a low degree of spatial variability across participants, so individual participant data were entered into a group-level t-test. ROIs were identified by thresholding group level t-maps at ($p < .005$).

The face-selective and body-selective ROIs were used as masks for further analyses using data from the main experiment. I wanted to see if various types of core and conceptual representation relied on the same neural systems as the visual processing of bodies and faces. Figure 18 shows that two body-selective regions (bilateral EBA) and one face-selective region (right posterior superior temporal sulcus) were preferentially involved in representing core self versus core representations for other people. The discussion section below provides a detailed treatment of this subject.

Controlling for Visual Luminance Effects

In Phase 1, the printed word stimuli differed in length. For example, “Acquaintance” is 12-letters-long, whereas “Relative” is 8-letters-long. Because stimuli consisted of white letters on a black background, differences in length translate to differences in luminance, which could have an effect on neural processing in the visual system. In contrasts between Phase 1 conditions, genuine core representation could be confounded with the effects of luminance. For Phase 2 contrasts, this was not an issue. Because the same set of properties was used in every Phase 2 condition, any effects related to luminance would drop out. Therefore, the following analysis pertains only to Phase 1 contrasts.

In order to account for luminance effects, I performed an amplitude-modulated regression to identify voxels in the brain whose BOLD signal correlates with luminance. I then compared these voxels with those significant in the experimental analyses described above.

A single amplitude-modulated regressor was created for every property stimulus presentation in the experiment. Phase 1 stimuli were not used because difference in length would covary with the cognitive effects of interest: Phase 2 stimuli were identical across conditions. For each trial, the amplitude of the convolution kernel was equal to the length of the property stimulus. For example, “tall” received a value of 4, whereas “confidence” received a value of 9. This regressor was entered into a GLM along with regressors of non-interest, such as the polynomial drift model, and fixed-amplitude regressors for each experimental condition (including properties). Because of the considerable power afforded by regressing on all complete trials, I thresholded the resulting activation at the group level at (p < .0001), a common threshold for functional localizers such as this (Kanwisher, McDermott, & Chun, 1997).

Supplementary Figure 1 through Supplementary Figure 8 present the results of this analysis. Voxels whose signal correlated with luminance are shown in blue, while voxel maps from other analyses are shown in red; the overlap is shown in yellow. The effect of luminance was significant in primary visual cortex, taking up most of BA 17, with limited encroachment beyond it. Supplementary Figure 1 shows the conjunctions for the four individual persons in Phase 1. Unsurprisingly, the overlap was greatest for the conditions with the most luminant stimuli: $Relative_1$, $Friend_1$, and $Acauaintance_1$. Supplementary Figure 2 shows the contrasts between these conditions. Again, large differences in stimulus length (e.g., between "Acquaintance" and "I") were associated with the greatest overlap with the luminance analysis. Supplementary Figure 3 shows that closeness was confounded with stimulus length, but that the effect was limited to BA 17. Supplementary Figures 4, 5, 6, 7, and 8 demonstrate that there was little overlap in the remaining analyses, because the stimuli were sufficiently low in luminance (for the conjunctions) or similar in luminance (for the contrasts).

Discussion

A central theme of this discussion is that person representation draws upon whatever information is available and relevant to the task at hand. Core and conceptual representations for the self, other persons, and groups of people require the coordinated efforts of multiple neurocognitive systems, such as those that underlie vision, action planning, and memory. Each of these contributes insofar as the information it provides is available and relevant for the type of representation being constructed. Following a brief overview of core and conceptual representation, this discussion approaches the results in two ways. First, I describe each of the relevant neurocognitive systems in detail, describing how each contributes to person representation. I will start with systems subserving core representation and finish with conceptual representation, but there is some overlap in the middle. Next, armed with a network level understanding of the results, I will revisit the three issues raised at the end of the introduction. Figure 17 presents this organizational scheme.

To begin my overall summary, I propose that core self is the unification of visual, sensorimotor, and spatial information that captures the basic characteristics of what it means to be a person. Core self is grounded in perception of one's visual appearance, ownership of one's body, agency for one's actions, and location in an embodied, egocentric frame of reference. When participants generated core self representations in the current experiment, their brains assembled the necessary information using the particular neural circuitry involved in processing each type of information. Together, the visual, premotor, lateral parietal, and medial parietal cortices constitute a neural system for generating core self-representations.

Core representations for other persons and for groups of people rely on similar processes as core self. To that extent, the self serves as a template for person representations in general. The differences between different types of core representation reflect differences in the availability and relevance of particular information streams. For example, body imagery situated in egocentric space is better suited for the self and other persons, whereas visual imagery situated in extrapersonal space is better suited for representing groups.

Conceptual representation follows a parallel logic: it too draws on the information that is available and relevant for the task at hand. Conceptual self, like core self, often makes use of visual, sensorimotor, and spatial information. However, conceptual self additionally requires the representation of one's personality traits, personal histories, and specific characteristics stored in semantic memory. This process requires maintenance and evaluation of conceptual information retrieved from semantic and episodic memory, and the attribution of these properties to oneself (Abu-Akel & Shamay-Tsoory, 2011).

Conceptual representations for other persons and for groups of people also use the self as a template. Just like with core representations, the differences between different types of conceptual representations reflect differences in the availability and relevance of particular information streams.

In the following sections I describe each of the neural systems involved in generating core and conceptual representations for the self, other persons, and groups of people (see Figure 17 for a list of neural systems). The particular functions that I ascribe to various brain regions and networks are well substantiated in the literature, but these are not the only possible interpretations consistent with the data (the "reverse inference problem"). For any given brain region, there are usually multiple functions ascribed to it in the literature. Here, I give plausible interpretations for the activations observed, but the evidence in this case is correlational, not causal. Nonetheless, the current experiment provides empirical evidence that person representation is grounded in neural systems whose more basic functions are well understood.

Visual Processing for Core Representations

Visual processing in occipitotemporal cortex plays an important role in both core and conceptual representation, as can be seen in almost every analysis I performed (Figure 23, all Tables). With respect to core representation, I interpret these findings according to three different functions of the visual system. Without being too cavalier in ascribing a one-to-one relationship between anatomical location and function, I suggest that the visual system plays three roles in my findings: perception, imagery, and the construction of object files. Of these, the latter two are directly involved in core representation.

Visual perception. The neural processes underlying low-level visual perception are relatively well understood (see Nassi & Calloway, 2009, for an excellent review). Primary visual cortex in BA 17 is principally concerned with the low-level perceptual features of a visual stimulus, such as color or the orientation of contours. From BA 17, visual processing diverges into two functionally dissociable but massively overlapping "streams": a ventral stream for processing object features and facilitating categorization, and a dorsal stream for processing motion and spatial information (Goodale & Milner, 1992; Mishkin & Ungerleider, 1982). Both streams include parts of BA 18 and 19, the next steps in the visual processing hierarchy after BA 17.

Beyond the occipital lobe, the ventral stream continues into the fusiform gyrus and other parts of the ventral temporal lobes where features of increasing complexity are extracted in service of recognizing objects, faces, and scenes (Stevens, Tessler, Peng, & Martin, 2015). The dorsal stream progresses into the dorsal temporal and parietal lobes where spatial data such as expansions and contractions, speed and direction, and translational motion help guide one's actions (Perry & Fallah, 2014).

Visual perception undoubtedly contributed to some, but not all, of the effects observed. In particular, overall luminance differed between some conditions, and between most conditions and the implicit baseline. These effects were present largely in BA 17, the primary visual processing region in which one would expect to see low-level perceptual effects (Pollen, 2011).

The confounding effects of luminance on various analyses are shown in Supplementary Figure 1 through Supplementary Figure 8. Voxels in blue are those whose activity correlated with phrase length, and therefore with luminance. Voxels in red are those that were significant for a given analysis, and voxels in yellow are the overlap between the two. As expected, the Phase 1 condition $Acquaintance_1$ appears to have generated the largest proportion of the overlap with areas associated with luminance. This is unsurprising, because "Acquaintance" was the longest word, and therefore the brightest stimulus, used in Phase 1. Crucially, visual activations of interest in BA 18 and 19, and beyond into fusiform gyrus, are largely unsullied.

Visual imagery*.* According to grounded theories of cognition (e.g., Barsalou, 2008; Harnad, 1990), neural systems for perception, action and interoception, such as the visual system described above, are recruited in service of higher-order cognitive functions. Perceptual representations are reinstantiated or "simulated" in service of some offline cognitive goal. For example, representing object categories activates brain areas responsible for representing the visual and motor properties of objects in those categories (Martin, 2007; Thompson-Schill, 2003). Visual imagery in particular has been shown to recruit many of the same brain regions as visual perception. One seminal study found shared activation in BA 18 and 19, as well as lateral and medial aspects of the parietal lobes, for both visual perception and imagery conditions (Kosslyn, Thompson, & Alpert, 1997). Some studies have even found imagery related activation in BA 17, although results are much more consistent in later visual areas (Kosslyn, Ganis, & Thompson, 2001).

By localizing regions of interest responsible for perceiving faces and bodies, and calculating activation in the main experiment within those ROIs, I demonstrated the role of visual imagery in core representations for the self and other persons. I localized three regions of interest for the visual perception of bodies: bilateral extrastriate body area (EBA) and right fusiform body area (FBA). These were consistent enough across participants that ROIs were defined at the group level. I also localized four regions of interest for faces: bilateral fusiform face area (FFA), and posterior superior temporal sulcus (pSTS). These were spatially inconsistent across participants, so ROIs were defined at the individual level. Statistics calculated within those ROIs were then analyzed at the group level (Figure 18).

One striking finding is that core self, but not core representations for other persons, recruits visual body imagery (Figure 18A). The EBA was activated during Phase 1 for self, but was not for other persons. First described as an occipitotemporal region involved in the visual processing of human bodies (Downing, Jiang, Shuman, & Kanwisher, 2001), the EBA has since been shown to be recruited by a diverse set of body-relevant cognition. For example, EBA has been implicated in observing one's agentic arm movements (Astafiev, Stanley, shsulam, & Corbetta, 2004), and even imagining one's location within one's body (Arzy, Thut, Mohr, Michel, & Blanke, 2006; Heydrich & Blanke, 2013). Crucially, EBA helps to distinguish between the self and other people, by distinguishing self versus another's bodies (Myers & Sowden, 2008; Saxe, Jamal, & Powell, 2005) and actions (David et al., 2007). I suggest that the EBA is part of a broader network for representing one's body (body ownership), grounding it in one's current location (first-person perspective), and recognizing one's actions as one's own (agency). The lateral parietal lobes, described below, are another crucial component of this network for core self.

The visual areas I found correlated with closeness in Phase 1 (Table 3), similar to those preferential for the self vs. others (Table 10), overlapped to a large extent with the EBA (Figure 19). The overlap was not perfect: the activation extended further ventrally, into other lateral occipitotemporal areas associated with action imitation and action concepts (Lingnau & Downing, 2015).

Results concerning core representation in face-selective regions are presented in Figure 18B. There was a preference for core self in right pSTS, an area involved in the visual processing of biological motion in addition to faces (Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004), but there was no straightforward pattern in the left pSTS, nor in either FFA. One parsimonious interpretation is that core representations involve bringing to mind information that is both available and relevant. For the self, bodily information is always available, because people are always present in their bodies. Similarly, simulation of facial appearance is an available substrate for representing both the self and other persons, because faces are crucial sources of information in everyday social interactions (Spencer, et al., 2011).

Considering specific individuals, rather than grouping all non-self persons together, each individual condition activated unique visual processing areas (Table 1 & 2). Especially prominent are the large areas unique to either I_1 *or Acquaintance*$_1$, but every individual has at least some unique activation. This may reflect differences in the availability of different types of visual information to be simulated, related to the variety of environments and social settings in which participants interact with different people. Moreover, interactions with acquaintances are more superficial than for friends and relatives, so the available and relevant information for them may be more visual in nature, as opposed to the types of information I will discuss next. Regardless, extrastriate and fusiform visual were widely activated for core representations of persons. It should be noted that the complete four-way conjunction might reflect the perceptual and motor requirements of the task, rather than core representation. The non-complete conjunctions reflect those aspects of core representation that are not shared across all persons, and thus are more likely to not contain task activations.

Object files. With respect to the visual processing hierarchy described above, object files are “mid-level” representations. They are post-perceptual, in that they go beyond individual features such as color and motion, but they are pre-conceptual, in that they do not rely on categorizations of the objects they represent (Perry & Fallah, 2014). Object files serve to disambiguate and individuate sources of information embedded in noisy environments. As such, they are a perfect vehicle for core representations for persons: they point to particular individuals, but are agnostic to higher-level conceptual information.

I propose that upon receiving an individual in Phase 1, participants generated an object file for that individual, bound to some minimal sensorimotor simulations (e.g., the face and body imagery just described, or the somatosensory or motor imagery I will describe below), but independent of the conceptual information that would soon be brought to bear in Phase 2. The current results suggest dorsal stream implementations of object files for three reasons. First, spatial information processed in the dorsal stream is thought to be one mechanism by which object files arise, individuating objects by their unique positions in mental space (Perry & Fallah, 2014). Second, neuroimaging data have implicated the posterior intraparietal sulcus, near the junction of the inferior and parietal lobules and occipital BA 19, in the representation of object files (Xu & Chun, 2009). Finally, my own data show increased activation at this boundary for $\textit{Individuals}_1 > \textit{Groups}_1$ (Table 5). I propose that core representations for individuals recruit these areas because individuals, by definition, can be individuated: they occupy discrete locations in space. Groups, on the other hand, cannot be localized in space, and therefore cannot be individuated in the same way. This is, of course, speculative: careful behavioral manipulations in future studies would be required to confirm the involvement of object files in constructing core representations for persons.

Interestingly, the semantic prompts activated these areas to an intermediate degree between individuals and groups (as can be inferred by their absence in Table 4, and from Supplementary Figure 9A). Perhaps, to a limited extent, the prompts “Physical” and “Genetic” instantiated person-related imagery as a framework for understanding what are otherwise rather abstract notions.

Premotor Cortex and Supplementary Motor Area in Core Representation

The premotor cortex has been implicated in body ownership (Ehrsson et al., 2005; Ehrsson, 2004), agency (Tsakiris, 2010), the sense of one’s own agency (Powell et al., 2010), and understanding the agentic actions of others (van Overwalle & Baetens, 2009). The SMA has likewise been implicated in agency (Moore, Ruge, Wenke, Rothwell, & Haggard, 2010; Nahab et al., 2011; Tsakiris et al., 2010). Specifically, the SMA likely provides a copy of efferent motor signals, generating the prediction against which sensory afference is compared to produce the sense of agency (Haggard & Whitford, 2004). In light of these findings and my previous results (Figure 3-7), I expected that these regions would be particularly instrumental in representing core self, and perhaps also in representing other persons and groups of people.

The *Individuals*$_1$ > *Groups*$_1$ contrast demonstrates that BA 6 (the union of premotor cortex and SMA) is preferential for persons over groups of people (Table 5, Figure 21). This makes sense: persons can be imagined to possess agency for their actions and ownership over their bodies, but groups of people cannot (mental states are perhaps a different matter, as I will discuss later). Surprisingly, these differences occurred below baseline, even for core self: no Phase 1 conditions show significant activation above baseline in BA 6, with the exception of *Friend*$_1$ (Table 1 & 2). In fact, a large cluster in primary somatosensory cortex for *Friend*$_1$ > *Relative*$_1$ indicates an even more embodied core representation for Friend. Likewise, Table 1 shows primary somatosensory and motor cortical activation for *Acquaintance*$_1$, but I found no other activations above baseline in these areas for persons.

Why are these important differences all occurring below baseline? One explanation could be my choice of baseline task: the button press response to the auditory tone may call upon the action planning functions of BA 6, overshadowing its role in generating core representations. Another explanation could be that BA 6 contributes to feelings of body ownership and agency online by integrating sensorimotor signals, as in the rubber hand illusion and the performance of intentional actions, but that integration of additional information is necessary in order to represent body ownership and agency. Tsakiris (2010) proposes that the output of BA 6 must be integrated with visuospatial, somatosensory, and proprioceptive signals in order to generate a more complete body model. He points to the role of lateral parietal cortices, to which I turn next.

Lateral Parietal Cortex and Core Representation

The lateral portion of each parietal lobe is comprised of a superior (SPL) and an inferior lobule (IPL), separated by the intraparietal sulcus (IPS). The IPL itself is comprised of the supramarginal and angular gyri (BA 40 & 39). The IPL, together with the posterior extremes of the superior temporal gyrus (pSTG) and sulcus (pSTS) informally comprises the *temporoparietal junction* (TPJ). There is no official standard defining the extent of the TPJ, but typically the term is used to label cortical territory close to the posterior extreme of the sylvian fissure. As one progresses further anteriorly and superiorly into the supramarginal gyrus, or inferiorly into the STG, the label of TPJ becomes less common. Here, I will use it only to describe temporal and parietal cortices in the vicinity of the posterior extreme of the sylvian fissure.

The right TPJ (particularly supramarginal gyrus) is active for I_1 and to a lesser extent $Friend_1$ (Table 1, Figure 20), implying a role in core representation for these individuals. Table 10 confirms that this activity is preferential for the self. Figure 20, and the bar graph in Supplementary Figure 10a, show just how much more active the right TPJ was for self vs. other people. Contrasts in Table 2 show that the distinction also applies to the left TPJ. In Tables 4 and 5, supramarginal activation is highly preferential for individuals over groups, although this activation is more dorsal than what is typically considered TPJ, and lies mostly in the IPS (Figure 21).

Why is rTPJ so intimately involved in core self, and why is rTPJ of little to no importance in representing other persons and groups of people? A likely answer is that right TPJ grounds representations of the body and of intentional actions in egocentric space, a property known as first-person perspective (1PP). Representations of other persons and groups of people are likely grounded in a different type of space, a process that recruits areas in the medial parietal lobes.

Uniquely situated to receive somatosensory, vestibular, and visual information needed to establish an egocentric coordinate system, the TPJ has been implicated in engaging in spatial, motor, and social tasks from a 1PP (Vogeley & Fink, 2003). Out-of-body experiences have been linked to the TPJ: psychophysical manipulations that produce illusions of 1PP activate the TPJ, electromagnetic stimulation of or damage to the TPJ produces illusions of 1PP (Blanke, 2012; Blanke & Metzinger, 2009).

Complementing its role in representing 1PP, the TPJ is a nexus for representing body ownership and agency (Ruby & Decety, 2001). As described in the introduction, the TPJ is implicated it in the false sense of body ownership in the rubber hand illusion (Tsakiris, 2010), the sense of agency (Farrer et al., 2008; Nahab et al., 2011, Tsakiris et al., 2010), and violations of expected agency (Haggard & Chambon, 2012).

Why then, is there TPJ activation for *Friend*$_1$ as well? Communication between the IPL and pSTS, together with premotor cortices, is thought to underlie action understanding in the so-called “mirror neuron system” (Urgen & Miller, 2015). If core representation depends on attributions of agency, one would predict the involvement of the TPJ. Indeed, the TPJ has been implicated in attributing agency to another person (Ruby & Decety, 2001). Perhaps the friends my participants imagined, being conspecifics, were prone to actions easily understood within a framework of the participants' own agency. This would explain both the TPJ and premotor activations seen for *Friend*$_1$.

For individuals more generally (*Individuals*$_1$), the considerable bilateral activations in supramarginal gyrus are more dorsal than what is typically included in the TPJ (Table 4). In fact, a large part of those activations, particularly when contrasted with *Groups*$_1$, resides deep in the IPS (Table 5 & Figure 21). Along with premotor cortex, IPS has been associated with illusory hand ownership in PET and fMRI studies, implying a role in body ownership (Blanke, 2012; Limanowski & Blankenberg, 2015), and in understanding the intentional actions of others (van Overwalle & Baetens, 2009). One study found IPS activation for an "agentic self" condition in which participants made volitional choices versus when the choice was made for them (Powell et al., 2010). That study cleverly controlled for the actual movements that participants made, meaning that the IPS activations were specific to the experienced sense of agency. Interestingly, my visual body processing localizer revealed a large swath of activation in the IPS and SPL, as has been previously observed (Weiner & Grill-Spector, 2011), further linking core representation with body imagery.

Precuneus, Posterior Cingulate Cortex, and Retrosplenial Cortex in Core and Conceptual Representations

My discussion of medial parietal cortex focuses on three posterior areas along the medial parietal wall: the precuneus (BA 7 & 31), the posterior cingulate cortex (PCC, BA 31 & 23), and the retrosplenial cortex (BA 23, 30, & 29). Note that the demarcations are approximate, and there is overlap between them. Because these three areas were involved in both core and conceptual representations, I will discuss both roles here. The anterior portion of the medial parietal lobes, consisting of somatosensory and paracentral areas, are functionally distinct and will not be discussed here.

The precuneus is involved in numerous self-referential activities (Cavanna & Trimble, 2006). Of particular relevance to core self is its role in generating an internal spatial model of the outside world from one's own perspective. That is, the precuneus is critical to process of imagining, in egocentric terms, the world and how one might interact with it in the present moment (Land, 2014; Mahayana, Tcheang, Chen, Juan, & Muggleton, 2014; Vogeley & Fink, 2003). Of particular relevance to conceptual self is its role in retrieval of autobiographical memories from a first-person perspective (Freton et al., 2014). Together, this is evidence for the precuneus playing a pivotal role in imagining oneself relating to one's world from both a synchronic and a diachronic perspective. As will be seen in a moment, both core and conceptual self recruit the precuneus.

The PCC is implicated in episodic memory retrieval, particularly memories involving close others (Maddock, Garrett, & Buonocore, 2001). Similarly, PCC represents the familiar places in which remembered events occurred (Sugiura, Shah, Zilles, & Fink, 2005). This is precisely the type of information that should constitute conceptual representations for other people: knowledge about a person's enduring traits based on memories accrued over time. Indeed, judgments regarding visual appearance for self, a close other, and an unfamiliar other all engage the PCC (Moran, Lee, & Gabrieli, 2010).

The adjacent retrosplenial cortex enables translation between egocentric and allocentric coordinates, particularly with respect to retrieved episodic memories (Vann, Aggleton, & Maguire, 2009). Drawing on scene-related visual information from the parahippocampal place area, the retrosplenial cortex facilitates spatial navigation and memory by situating a scene in the context of a broader environment. For example, when recognizing a picture of your kitchen, the retrosplenial cortex would place it in the context of your house, an environment for which you possess rich episodic memories (Epstein, 2008).

Medial parietal cortex in general is crucial for generating vivid imagery, often from episodic memory, and situating the imagined people and things in a mental spatial model of the outside world. Whereas precuneus activity tends to be more self-related and egocentric (Cavanna & Trimble, 2006), activity in the PCC and retrosplenial cortex tends to be more other-related and extra-personal (Vann et al., 2009). I do not suggest that this is a clean double dissociation, but it provides a useful heuristic for interpreting my findings.

Results indicate that the role of medial parietal cortex in core representation is complex. All four of the person conditions activated the precuneus, but only $\textit{Friend}_1$ and $\textit{Acquaintance}_1$ activated the nearby PCC (Table 1). Comparing core self with all three non-self others (Table 10) further suggests a special role for the PCC in core representations for other persons. However, in the specific contrasts between persons (Table 2), the PCC pattern is only borne out in $\textit{Friend}_1 > I_1$. Therefore, the nuances with respect to particular individuals should be interpreted with caution. A multivariate analysis, which is not possible with the current design, would likely be informative.

The precuneus activation is particularly interesting, because although a large proportion of voxels in the precuneus were activated by at least one individual, only nine voxels were significant for all four (Figure 22). This means that precuneus was representing all of the individuals, but in slightly different ways. The precuneus is sensitive to observation of social interactions between others, but the nuances of this role have yet to be elucidated (Petrini, Piwek, Crabbe, Pollick, & Garrod, 2014).

For conceptual representations, the same general pattern persists in Phase 2: precuneus for conceptual representations for self, PCC and retrosplenial cortex for others (Table 7). Here is another indication that core versus conceptual is not an absolute distinction: the medial parietal lobe behaves quite similarly for both, at least at the level of specific persons. In a moment, I will describe subtle differences at the collapsed task level of individual persons and groups of people more generally.

Medial parietal cortex appears to be critical for core representations of both individuals and groups. During Phase 1, I observed robust activation in left PCC for *Groups*$_1$ (Table 4), and in left retrosplenial cortex for *Groups*$_1$ vs. *Individuals*$_1$ (Table 5). Given a general pattern of less activation for groups than for individuals (Table 5), it is noteworthy to see the opposite in retrosplenial cortex (Figure 22). I interpret this finding in light of my central premise that person representations draw upon whatever sources of information are relevant and available. For example, the lateral parietal and premotor network for body ownership and agency is of no use because groups have neither bodies nor individual agency. As with non-self persons, it appears that core representations for groups consist of vivid visual imagery situated in spatial contexts drawn from memory. For groups, these representations are especially important because embodied information is not available.

Oddly, the medial parietal activations for groups disappear in Phase 2. Betas in PCC and retrosplenial cortex do trend positive, but they do not meet my significance criteria. Perhaps, as specific properties are being predicated of them, situated visual imagery gives way to other sources of information (Table 8), such as embodiment (SMA and left premotor cortex), and abstract semantic knowledge (middle temporal gyrus: see below for discussion).

In the medial parietal lobes, the distinction between core and conceptual begins to blur: the process of generating an internal model of the outside world is equally important for both. The differences, if there are any, may lie in the temporal perspective taken. Core representations require a synchronic focus, wherein one imagines the world, the self, and other people as they are; conceptual representations additionally draw on diachronic information derived from episodic memory. In each case, whereas self is grounded in egocentric space by virtue of the lateral parietal cortices and precuneus, my findings suggest that representations for other persons and groups of people are situated in an extra-personal mental space by virtue of activity in the PCC and retrosplenial cortex.

Visual Processing and Conceptual Representations

During Phase 2, visual activation was more prominent relative to baseline that it was during Phase 1 (Figure 23), but the differences between conditions had nearly disappeared (Tables 7 & 9). Two observations are worth noting. First, except for a few voxels in $Acquaintance_2 > I_2$, BA 17 was not differentially activated for any of the contrasts. This was to be expected if BA 17 is only processing visual perception, because Phase 2 stimuli were controlled for luminance across conditions. Compare this observation to Phase 1 (Table 2), in which more-luminant individual stimuli (e.g., "Acquaintance") activated BA 17 to a greater degree than less-luminant stimuli (e.g., "I").

Second, at the collapsed task level, all three tasks activated a large portion of the ventral visual stream, especially in the left hemisphere (Table 8). This suggests that the conceptual knowledge being brought to bear during all three tasks is, to a large extent, represented visually. Importantly, the differences between tasks are null, implying that visual content is contributing equally across tasks (Table 9). Again, however, the possibility remains that activations seen in every condition might reflect task-based activity; only the non-complete conjunctions can be more unambiguously interpreted as reflecting self representation.

The ventral temporal lobes contribute visual content to a broader left lateralized semantic network including the lateral temporal cortex and ventrolateral prefrontal cortex (Thompson-Schill, 2003). Below, I will discuss this network in more detail with respect to those regions, and see where the differences between individuals and groups lie.

Premotor Cortex and Supplementary Motor Area in Conceptual Representation

During Phase 2, the premotor cortices, including SMA, play a much more prominent role than they did in Phase 1. SMA and left premotor cortex are active for all four individual persons (Table 6) and for all three tasks (Table 8). Right premotor cortex is active for *Individuals*$_2$ and *Prompts*$_2$ but not for *Groups*$_2$. The same right premotor cluster is active at ($p < .005$) for *I*$_2$, *Relative*$_2$, and *Friend*$_2$, but the cluster is too small (28 voxels) to appear in the conjunction. Again, activity that is uniform across conditions should be interpreted with caution, because it may merely reflect the particulars of the task. Interestingly, contrasts between persons (Table 7) and between tasks (Table 9) reveal no differences in premotor cortex and SMA, except for *Prompts*$_2$ > *Groups*$_2$.

The activations in the left premotor and motor systems in Phase 2 (Table 6 and Table 8) may in part reflect the button place that occurred in response to the properties. I consider this unlikely to be a substantial confound because there was also a button press during the active baseline. Moreover, any effects of the button press should be cancelled out in contrasts between conditions involving a button press response. However, the potential confound should be recognized with respect to the conjunctions.

As discussed earlier, the premotor cortex and SMA play a significant role in body ownership and agency. The current results suggest that the relevance of premotor cortex and SMA are not limited to core representation. Two interpretations are consistent with this result. On the one hand, it could be that these regions are representing the embodied aspects of the properties. Many of the properties were chosen to be physical in nature, some of them specifically pertaining to sensorimotor traits (Figure 12). On the other hand, it could be that core representations, insofar as they rely on body ownership and agency and are processed in premotor cortex and SMA, remain active during Phase 2. This makes sense for two reasons, one empirical and one theoretical. First, this is exactly what I observed in my previous experiment (Figure 5). Second, the predication process central to the individuals task and the groups task requires a subject: the relevant core representation is maintained online so that the property may be predicated upon it. That does not, however, explain the *Prompts*$_2$ activations, which should not have involved predication upon a person. Further study is required to answer this final point.

These interpretations are not mutually exclusive: they can both be sound, provided it is granted that core and conceptual representation are not entirely distinct phenomena. The information processed in premotor cortex and the SMA is brought to bear whenever it is available and relevant, regardless of whether it is in service of core or conceptual representation.

Medial Prefrontal Cortex and Conceptual Representation

The medial prefrontal cortex (mPFC) is the common denominator of neuroimaging studies pertaining to the self in particular and social cognition in general. Activity in the mPFC is associated with such diverse functions as reflecting on one's traits, dispositions, attitudes, mental and emotional states, and imagining the beliefs, thoughts, feelings, and desires of other people (Mitchell, 2009).

Accordingly, I found mPFC to be associated with conceptual representations for persons (Table 6). Again, I did not find mPFC involvement in core representation for persons (Table 1), replicating a key finding of the previous experiment (Drucker et al., submitted; Figure 3). During Phase 2, I observed activation primarily for the self in the dorsal aspect (dmPFC, roughly medial BA 8 & 9), and for other persons in the ventral aspect (vmPFC, roughly medial BA 10 & 11) of the mPFC (Table 6 & 7, Figure 24). As I will discuss in a moment, this goes against conventional wisdom, but is not entirely surprising. At the collapsed level of analysis, overwhelmingly, I observed mPFC activation during Phase 2 for individuals, not for groups or during the semantic task (Table 8 & 9, Figure 24). Interestingly, there was extensive mPFC activation for the semantic task during Phase 1 (Table 4). This unexpected observation is consistent with multiple explanations; perhaps participants directed less stimulus-directed attention to the prompts, failing to deactivate the default-mode network. Another possibility is that participants processed the prompts ("Physical" and "Genetic") with respect to imagined persons, despite instructions designed to discourage person representation.

Medial prefrontal cortex activation in conceptual representations for the self and other persons.

The vmPFC plays a key role in representing the conceptual self, a finding corroborated ubiquitously from early empirical studies (Kelley et al., 2002) to more recent meta-analyses (Denny, Kober, Wager, & Ochsner, 2012; Northoff et al., 2006; van Overwalle, 2009). The centrality of vmPFC to self processing is likely related to emotion processing: the vmPFC is densely interconnected with cortical and subcortical networks for processing internal states and the feelings that accompany them (Abu-Akel & Shamay-Tsoory, 2011).

The dmPFC is involved in mentalizing about others, inferring their beliefs, desires, and intentions (Isoda & Noritake, 2013). Accuracy in inferring others' preferences is associated with greater activation in the dmPFC, as well as increased functional connectivity between the dmPFC and both the TPJ and PCC/precuneus, two other classic self areas also implicated in theory of mind. In the same study, using one's own preferences as a reference point was associated with activation in vmPFC (Kang, Lee, Sul, & Kim, 2013).

In general, the dmPFC is thought to preferentially represent other persons, whereas the vmPFC represents the self (Wagner et al., 2012). A meta-analysis of fMRI experiments studying conceptual self and others found a spatial gradient in medial prefrontal cortex, where others were represented more dorsally and the self was represented more ventrally (Denny et al., 2012). Therefore, it is somewhat surprising that I found the opposite pattern (Table 11, Figure 24).

As a possible explanation, the dmPFC, more than the vmPFC, is associated with cognitive judgments pertaining to self-related stimuli (Abu-Akel & Shamay-Tsoory, 2011; Northoff & Bermpohl, 2004). Therefore, the dmPFC activation I observe for self-related conceptual representation may be due to a depth-of-processing effect: because participants had so much more prior knowledge with respect to themselves than for other persons, they had more information to consider when making their judgments.

Conversely, the vmPFC activation observed may reflect the use of the conceptual self as a template for close others: mentalizing about the self and close others activates overlapping medial prefrontal regions (Heatherton, 2011; Krienen, Tu, & Buckner, 2010). In the current experiment, I showed that medial prefrontal activation for conceptual representations correlated with closeness (Table 3), suggesting that neural overlap between self and other is related to closeness. By contrast, activation in vmPFC is less active for a distant other, such as George W. Bush (Powell et al., 2010) or Oprah Winfrey (Drucker et al., submitted). If the role of the vmPFC in self processing is related to its role in emotion processing, then it is hardly a surprise that it would be involved in representing close others in addition to the self.

Regardless, the dorsal/ventral gradient observed in conceptual representations for others and the self is a trend, not the law: much more consistent is the finding that both parts of the medial PFC are involved in representing mental states for both self and other, with a general preference for self (Jenkins & Mitchell, 2011). Again, this preference is inconsistent: which parts of the mPFC are active for self or other persons varies substantially between studies (Denny et al., 2012; Legrand & Ruby, 2009; Mitchell, 2009; Vanderwal et al., 2008).

Medial prefrontal cortex activation in conceptual representations for individual persons and groups of people*.* *I* observed an extremely robust pattern of dmPFC and vmPFC activation for individuals (Figure 24), but not for groups or the semantic task, during Phase 2 (Table 8). I conclude that the mPFC plays a unique role in generating conceptual representations for individuals in possession of a mind. Whereas this is precisely the pattern I expected for individuals, I was somewhat surprised not to find a similar pattern for groups. For example, Contreras and colleagues (2013) found similar activation throughout the ToM network for the mental states of both groups and individuals. Interestingly, they found that multivoxel patterns within many of these areas distinguished between groups and individuals. Unfortunately, my rapid event-related design with catch trials does not lend itself to such an analysis. However, at an extremely liberal threshold ($p < .04$ voxelwise), some activation is observable for $\textit{Groups}_2 > \textit{Prompts}_2$. The spatial pattern, with peaks in dmPFC, PCC, and rTPJ, suggests that this is not coincidental. If participants were representing the collective minds of adults and children, such representations were not robust.

Lateral Parietal Cortex and Conceptual Representation

The left TPJ is highly preferential for conceptual representations for self versus other persons (Table 9, Table 11) although this activation is below baseline, and therefore absent from the conjunctions. The activation is extensive, ranging from the dorsal part of the IPL down through the TPJ and into the superior and medial temporal gyri (Figure 25). This pattern is borne out in several of my analyses, in subtly different ways. In Table 3, the IOS regressor captured a gradient in left TPJ from self to other according to the closeness of the individual. In Table 2 the same area, unsurprisingly, appears in many of the contrasts between individuals.

At the collapsed task level, lateral parietal activation is conspicuously absent from the remaining Phase 2 analyses (Tables 8 & 9). This is surprising: I expected that TPJ would play a pivotal role in conceptual representations for individual persons generally, and perhaps for groups as well, especially compared with the semantic task.

The TPJ is typically associated with mentalizing, that is, representing mental states, for the self and others (Abu-Akel & Shamay-Tsoory, 2011; Lombardo et al., 2010). The primary support for this conclusion comes from tasks in which participants must attribute beliefs that differ from their own to another person. In other words, these are tasks that probe theory of mind. ToM tasks have been repeatedly shown to recruit the TPJ (Abu-Akel & Shamay-Tsoory, 2011; Saxe & Kanwisher, 2003; Saxe, Moran, Scholz, & Gabrieli, 2006). I anticipated that conceptual representations, especially for mental properties, would recruit the TPJ. I was only partly right, in that left TPJ was preferentially activated by conceptual self versus other persons (Table 11, Figure 25).

Despite its notoriety for its role in mentalizing, the TPJ subserves diverse functions. I have already reviewed its role in various facets of core self: first-person perspective, body ownership, and agency. Another, perhaps more fundamental function of the TPJ is orienting attention. One early study identified TPJ activation as participants oriented their attention to the location of a target stimulus (Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000). A later study showed the TPJ is also involved in attentional capture by salient stimuli (Serences et al., 2005). Crucially, a more recent study directly compared attention attentional versus mentalizing functions of the TPJ, and found that they recruited the same voxels in the TPJ (Mitchell, 2008).

I propose that the various proclivities of the TPJ (e.g., body ownership, agency, 1PP, and representing mental states) are in fact manifestations of a single function: orienting one's perspective to a location in space. In terms of the examples just given, that location may be within one's body or mind, within another person's body or mind, or even in a fake body with proper induction (Blanke, 2012). My Phase 1 instruction, to think deeply about the individual being presented, induced participants to orient to the target person's mental perspective. The data suggest that this was most feasible for the self and for a friend (Table 1). The property-verification tasks in Phase 2, however, did not require participants to take on the perspective of other people: merely to describe them. As an aside, my baseline task required the orienting of attention to a highly salient auditory stimulus. This undoubtedly interfered with activation in the TPJ, and probably in other attention-related areas as well. As a consequence, differences between Phase 2 conditions could be observed, but overall levels of activation with respect to baseline were all negative in the TPJ. Future studies must consider the advantages and disadvantages of an active baseline.

Consequently, I suggest recategorizing mental state attribution as a core self representational process. In hindsight, this makes perfect sense. Mental states are, by definition, synchronic: they refer to the state of one's mind in the present moment. The more diachronic aspects of ones mental life, that is, personality traits, are what constitute conceptual self. As seen earlier, representing these enduring mental properties is the purview of cortical midline structures in the frontal and parietal lobes. Accordingly, one experiment used a factorial design to directly compare the neural correlates of understanding the goals underlying another person's behavior (i.e., a mental state) versus understanding that person's traits. Their findings for the goal/trait distinction map remarkably well onto my core/conceptual distinction. They found that trait attributions were associated with activity in the SMA and dmPFC, whereas understanding goals was associated with TPJ, premotor cortex, pSTS, and intraparietal sulcus (Ma, Vanderckhove, Van Hoeck, & van Overwalle, 2012).

Left Lateral Temporal Cortex and Conceptual Representations

The left lateral temporal lobe, particularly the middle temporal gyrus (MTG), has long been known to subserve the long term storage and retrieval of semantic knowledge (Thompson-Schill, 2003). It therefore comes as no surprise that the left MTG is activated during Phase 2 for every condition (Tables 6 & 8), because all three tasks involve the retrieval, representation, and evaluation of semantic knowledge.

The anterior portion of the left temporal lobe (BA 38, called ATL for short, or the temporal pole) is specifically involved in representing semantic knowledge with respect to persons. According to the "semantic hub" theory, the ATL is the hub of a semantic network that increases in abstraction along the posterior-anterior axis of the temporal lobe. The ATL, being the apex of this conceptual hierarchy, integrates lower-level sensorimotor features into higher-level amodal category representations (Hoffman, Evans, & Lambon Ralph, 2014). An alternative theory posits that the ATL specifically stores person knowledge underlying social cognition (Simmons, Reddish, Bellgowan, & Martin, 2010).

My results support the latter theory. Every Phase 2 condition activated the MTG, but only *Individuals*$_2$ activated the entire length of the MTG, from the posterior portion all the way into the ATL (Figure 26). Contrasts with *Groups*$_2$ and *Prompts*$_2$ both reveal activation in the ATL. The contrast *Individuals*$_2$ > *Prompts*$_2$ is particularly important (Figure 26), because it directly tests the hypothesis that person knowledge is privileged in the left ATL. Both tasks required activating the same semantic concepts, because each of the properties appeared exactly once in each of the three tasks. However, the individuals task required that those properties be activated with respect to persons, whereas the semantic task did not.

It is possible that activations preferential for individual persons in left ATL reflect the representation of persons' names: unlike the groups or prompts, the individuals are real persons with their own names. Support for this idea comes from the observation that lesions in the left ATL lead to specific naming deficits (Drane et al., 2008). If this were the case, it would have been expected that the ATL would have been activated in Phase 1, not just Phase 2, but the possibility should not be discounted. Moreover, representing an individual's name is itself semantic knowledge, and so an effect of naming would not be inconsistent with the above interpretations.

In short, the left lateral temporal lobe is involved more or less equally in representing conceptual knowledge for the self, other persons, and groups of people, but the ATL is preferential for individual persons. There were no substantial differences between persons, except for a slight preference for *Acquaintance*$_2$ (Table 2).

Lateral Prefrontal Cortex and Conceptual Representation

Left ventrolateral prefrontal cortex (vlPFC, equivalent to the inferior frontal gyrus), enables the retrieval of information from memory, particularly when such retrieval requires effort (Wagner, Pare-Blagoev, Clark, & Poldrack, 2001).

Specifically, pars orbitalis of the inferior frontal gyrus (BA 47, also called lateral orbitofrontal cortex) guides controlled retrieval from long term information stores in lateral temporal cortex; pars triangularis (BA 45) selects relevant items from among the retrieved representations; and pars opercularis (BA 44) selects an appropriate response (Badre & Wagner, 2007). It is therefore unsurprising that left vlPFC was activated during Phase 2, when conceptual information was brought to bear for all four individuals and for all three tasks (Tables 6 & 8, Figure 26).

By contrast, right vlPFC, especially pars opercularis (BA 44) has been implicated in retrieving perceptual information from episodic memory (Dobbins & Wagner, 2005). Interestingly, BA 44 in the right vlPFC was preferentially activated for properties associated with prompts, although this activation also was not significantly above baseline (Table 9).

Left dorsolateral prefrontal cortex (dlPFC, roughly equivalent to the middle frontal gyrus, or BA 8, 9, and 46 on the lateral surface of the brain) is involved in a wide range of effortful cognitive tasks, such as maintenance and manipulation of items in working memory, maintenance of goals during a task, and making decisions and constructing action plans in service of those goals (Fuster, 2001; Miller & Cohen, 2001). Specific to person representation, the dlPFC subserves the cognitive aspects of theory of mind, such as attributing mental states to a particular individual (Abu-Akel & Shamay-Tsoory, 2011; Carluer et al., 2015). As with vlPFC, dlPFC was involved in every task (Tables 1 & 8), because every condition required the maintenance of conceptual information and a decision requiring attentional resources. Although it was activated uniformly across individuals, dlPFC was preferentially activated during Phase 2 of the semantic task, perhaps due to difficulty of the task as indicated by slightly longer response times (Table 9, Figure 27).

A Process Model for the Experimental Task

The interpretations delineated above rely upon a set of assumptions: that participants are performing the task as instructed, that the task elicits core and conceptual representations as intended, and that all or most of the activations I observed reflected the constructs of interest and not some incidental effect of the task. I now briefly present a process model of how participants performed the task in the scanner, assessing the three issues just mentioned, and assessing my confidence that each is addressed.

For Phase 1, participants were instructed to think deeply about the individuals, group, or semantic prompt as soon as it was presented. Participants were explicitly given flexibility to interpret that instruction as they saw fit, but were implicitly guided towards an experiential, embodied perspective. A priori, there was no guarantee that Phase 1 would elicit core representations, in the current experiment or my previous experiment (Drucker et al., submitted). As a manipulation check, participants in both experiments were asked before and during the session to affirm that they understood the task, and after the session to affirm that they were thinking deeply about the individual, group, or prompt in Phase 1.

Additional evidence for the validity of the task comes from the brain networks that I identified with respect to Phase 1. The involvement of areas associated with body ownership, agency, the first-person perspective, and mid-level perceptual features support the claim that core representations were instantiated for individuals. Differences between individuals, such as greater involvement of TPJ for self and PCC for other, indicate a shift in perspective from peripersonal to allocentic space. For groups, there was a also a shift away from embodiment and towards visual imagery situated in allocentric space, as might be expected. Therefore, I conclude with a reasonable degree of confidence that during Phase 1, core representations were engendered, devoid of the reflective conceptual knowledge pertaining to diachronic traits, social roles, and so on that would follow in Phase 2. For semantic prompts, the recruitment of subcortical and medial prefrontal systems is puzzling (Table 4 & Table 5). I speculated earlier as to why that might be, but I hesitate to draw firm conclusions.

For Phase 2, participants were instructed to judge how well the property applied to the individual, group, or prompt. The expectation was that this task would require participants to generate rich conceptual representations for the individual or group previously presented, which would remain active. For the semantic task, only the property itself would be represented. In line with previous literature, the involvement of medial prefrontal and anterior temporal areas during the individuals task suggests that the manipulation was effective

During Phase 2, I suggest that four cognitive processes occurred. First, core representations from Phase 1 (for individuals and groups) were maintained: in the relevant brain networks, activation was maintained or even increased. Second, conceptual representations were generated in, sensory, motor, and lateral temporal systems, in order to answer the question “to what extent does the property apply?”

to the individual or group. For the semantic task, only the meaning of the property would have been represented. Third, the property was applied to the individual, group, or prompt, and a decision was made, primarily involving lateral prefrontal cortex. Finally, based on the selection, a button press response was made.

Addressing the Issues that Motivated the Current Experiment

The current experiment was designed to address three outstanding issues from a prior experiment (Drucker et al., submitted). I now interpret my current results with respect to these three issues.

Issue #1: Does the self enjoy a privileged status in the human brain, or is it a special case of more general processes? It has been argued that representing the self and other persons are in principle the same process; they differ according to which sources of information are relevant to each. For example, self-representations would be more interoceptive and embodied, whereas other representations would be more exteroceptive and perceptual in nature (Colzato, van den Wildenberg, & Hommel, 2013; Lombardo & Baron-Cohen, 2011). My findings are in accordance with this view: I found preferences for self versus others in EBA and TPJ, and for others versus self in medial parietal areas associated with visual imagery retrieved from memory.

I further suggest that information about the self may be used as a template for representing other persons, a position supported by overlap between self and other in both cognitive and neural domains. Early evidence came from the observation that the self-reference effect, a memory advantage when items are related to the self at encoding, is also seen for intimate others (Symons & Johnson, 1997). A more recent study, involving the developer of the IOS, had participants rate traits with respect to the self and other persons. A surprise source memory task (“who [sic] was each trait rated for?”) generated more confusions between self and close others than between self and non-close others (Mashek, Aron, & Boncimino, 2003). Similarity (sharing features) and familiarity (knowing who the person is) did not predict errors. The authors suggest that the knowledge structures underlying self and close others are overlapping, and that that in fact is what it means to be close. Accordingly, I found that closeness correlated with BOLD activation in several key core self areas in the ventral visual stream and the striatum (Table 3).

In terms of neural processing, the extent to which core self representations can be used as a template for other persons may also be related to the extent to which one can imagine herself in the others’ situation. Earlier, I described a unique case study in which a congenital amputee, named DD born without arms or legs, watched videos of people performing actions she either could or could not perform herself (Aziz-Zadeh et al., 2012). Observation of possible actions recruited premotor cortex and the TPJ, brain areas that are involved in body ownership and agency, and that I found relevant for core representation for persons. Observing impossible actions recruited these areas plus medial PFC and precuneus, brain areas that are involved in mentalizing, and that I found relevant for conceptual representations for persons.

These findings are in line with my central thesis that person representations draw on whatever information is relevant and available. Patient DD represented the bodies and actions of other people using the systems underlying her own body ownership and agency, but only when such information was available to her. Otherwise, she was obliged to use more abstract conceptual representations of the actions being performed. Incidentally, this may help explain the medial PFC activations I observed for the semantic prompts. Participants were trying to represent the concepts “physical” and “genetic” without knowing to whom they referred, like a predicate without a subject, or like patient DD contemplating actions she could not perform.

As for conceptual representations, it is thought that closeness moderates the degree of overlap between self and other, especially in the mPFC (Mitchell, Macrae, & Banaji, 2006; Wagner et al., 2012). Accordingly, I observed BOLD activation positively correlated with closeness in the left TPJ and dmPFC, and negatively correlated with closeness in the PCC and retrosplenial cortex (Table 3), two areas critically but differently important for conceptual person representation. I may interpret this phenomenon as metaphorical thinking: the self is a template for others, or conversely self-reflection is other-perception turned inwards. Or, I may interpret my findings in terms of overlapping knowledge structures (Mashek et al., 2003). I propose that the point is moot, because all of these perspectives are correct, and may be reconciled by a single overarching neurocognitive principle: the brain makes optimal use of its limited resources by recycling them, deploying them in diverse cognitive tasks whenever they are available and provide relevant information (Anderson, 2010).

Issue #2: Do core representations for groups exist? If so, are they instantiated by the same network as core representations for individuals? A general theme of my findings is that self-representation can serve as a template for other persons, but that groups of people do not elicit rich representations. For conceptual representations for groups of people, I observed activation in visual, premotor, prefrontal, and temporal cortices, but none of these areas were preferential or unique for groups versus individuals. The areas present indicate the retrieval of semantic information (vlPFC & MTG), low-level sensorimotor imagery (primary somatosensory, premotor, and motor cortices, SMA, extrastriate and fusiform cortices), and cognitive judgments made in service of the task (dlPFC). However, activations were not observed in the ToM areas active for individuals (mPFC, medial parietal cortex, IPL & TPJ), implying that participants were not deeply mentalizing about groups.

Likewise, behaviorally, participants rated properties-for-groups with the noncommittal “applies somewhat” option more often than “applies well” or “applies poorly”, especially compared with individuals or semantic prompts (Figure 28). Therefore, I conclude that conceptual representations exist for groups, but that they are richer, with more abstract mental content, for individuals. This runs counter to the conclusions of an earlier study, although that study used a very different contrast to identify activation related to mentalizing about groups (Contreras et al., 2013). Why did I find conceptual representations for groups to be so limited?

Scheepers and colleagues (2013) suggest that metaphorical extension of self-representations to groups may depend on identification with that group. That is, identifying with a particular group serves as a bridge, linking the neural representations for the self and for that group. Adults and children are somewhat ambiguous, and my population was mostly college students, precariously perched between childhood and adulthood. Future studies should explore the neural correlates of in-groups and out-groups, and the extent to which participants identify with them.

Issue #3: Does activity in the classic self network reflect the predication of properties upon people, or the representation of those properties per se? Some suggest that the vmPFC, dmPFC, PCC, and TPJ is a domain-general system for monitoring and modulating other cognitive subsystems, for example guiding attention to relevant persons and conceptual information (Dosenbach et al., Kang, Lee, Sul, & Kim, 2013). In this view, social cognition is just one domain in which these areas play a central role (Isoda & Noritake, 2013; Legrand & Ruby, 2009). Indeed, the DMN is known to be central to remembering the past and imagining the future, and DMN dysfunction is associated with pathologies whose effects go beyond social cognition (Buckner, Andrews-Hanna, & Schacter, 2008).

Conversely, it has been suggested that the classic self network is specialized for processing person-relevant information (Saxe et al., 2004; van Overwalle, 2009). Even the non-social functions mentioned above can be thought of as mentally "projecting" oneself into the past, future, and hypothetical situations (Buckner & Carroll, 2007; de Brigard, Spreng, Mitchell, & Schacter, 2015). In fact, one recent study experimentally disrupted DMN activity during rest using low-frequency rTMS and found an effect on self-awareness (Gruberger et al., 2015).

If this network is specialized for social cognition, then information encoded there should be specifically related to representing persons. Along these lines, a study using adaptation fMRI identified a region in vmPFC sensitive to particular personality traits. A separate vmPFC region, along with the precuneus, was sensitive to a trait inference process more generally (Ma et al., 2014). A meta-analysis found a bias in mPFC for mentalizing processes over more general reasoning, implying a domain-specific role for the mPFC in representing minds (van Overwalle, 2009).

The current findings support the latter interpretation. In particular, the contrast $\mathit{Individuals}_2 > \mathit{Prompts}_2$ reveals that cortical midline structures in the frontal and parietal lobes, and the anterior portion of the left lateral temporal lobe, are specifically generating conceptual representations for persons (Table 9, Figure 24 & 26). The same property stimuli were used in both tasks, and both required a judgment regarding the property presented on a given trial: the only difference was whether or not that property was predicated of an individual.

The current results indicate that conceptual representation, just like core representation, draws upon whatever information is available and relevant to the task at hand. Some of these information streams were shared among the three tasks. First, modality-specific systems processed the low-level features of the properties: ventral occipitotemporal cortex for visual information, premotor and supplementary motor cortices for embodied information, and lateral parietal cortex for spatial information. Second, left posterior MTG, a classic semantic knowledge area, processed the properties with respect to their categorical structure.

However, a separate network seems to be processing information specifically with regard to individuals. The medial prefrontal cortex represents a person's enduring personality traits and mental dispositions. The left anterior temporal lobe represents the more abstract, semantic knowledge associated with that person. Then, the medial parietal lobes aid in retrieving all of the above information from episodic memory, and in situating it within a spatial context. This last step also applies to groups: although they cannot be individuated, they can still be imagined in an extra-personal space.

Finally, across all tasks, the lateral prefrontal cortex, instrumental in reinstantiating all of this information in the first place, applies it to the appropriate individual, group, or prompt, and facilitates a decision.

Limitations

Time-invariant contributions to core self. I employed a task-based paradigm, in which participants generated core and conceptual representations for themselves, other persons, and groups of people. However, because core self is computed anew in each moment, some aspects of core self are likely always present, and therefore impermeable to the methods used here. The dependent variable in the current experiment is the timecourse of percent signal change in the BOLD signal for each voxel. As such, it requires that the neural correlates of the phenomenon under consideration vary with respect to time. For all of the brain regions that I have described as being involved in core representation, I successfully introduced temporal variance by manipulating the independent variables, that is, by drawing attention to core versus conceptual representations, and by using different subjects (e.g., “I”, “Acquaintance”, “Children”, etc.). Insofar as the core self is invariant with respect to time, I undoubtedly missed out on interesting neural activity. The neural systems involved in generating core self that are most likely invariant with respect to time are those that continually monitor homeostatic states, such as the brainstem, hypothalamus, and insula (Damasio, 1999).

Controlling for luminance. I did not adequately control for low-level visual properties of the stimuli. Words with more letters produced more luminance, which in turn varied with activation in participants’ early visual cortex. I observed luminance effects confounded with I_1 (by far the shortest word or phrase) in the BOLD response in primary visual cortex and the lateral geniculate nucleus of the thalamus. Follow-up studies should use filler symbols within the bounding box, such as dashes or asterisks, to balance luminance across all conditions.

Still, three inferences help us interpret the current data with respect to visual processing areas. First, my analysis of word length helps control for variance associated with luminance, indicating that BA 17 activity is primarily associated with low-level perceptual properties. Second, activation in BA 18 and beyond that is greater for less luminant stimuli (e.g., I_1 versus other individuals) is strong evidence for visual imagery. Third, all contrasts between Phase 2 conditions should cancel out any perceptual effects, because the properties used were identical in every condition

Active vs. passive baseline. Each has its advantages and disadvantages. Essentially, an optimal baseline is one that is as different as possible from the cognitive process of interest. In the case of the current experiment, trade-offs were inevitable, because core and conceptual representation draws on so many diverse cognitive functions. Here, I have laid a broad foundation: future experiments should focus on more restricted domains, with baseline tasks chosen accordingly.

Insula conspicuously absent. The insula, especially the right insula, is implicated in both body ownership and agency, both in lesion studies with patients and imaging studies with healthy adults (Karnath & Baier, 2010). Several psychiatric and neurological disorders of core self implicate the insula, including depersonalization and somatoparaphrenia (Sierra & David, 2011; Vallar & Ronchi, 2009). This finding is complemented by imaging studies that find insula activation correlating with RHI and other measures of body ownership and agency (Tsakiris, 2010).

The current results do not provide strong evidence for a role of the insula in core self representation. I suggest that this is due to the top-down nature of the task: participants were required to generate core representations for the self, other persons, and groups of people in response to a linguistic cue. There are advantages to my approach, in that I was able to observe the offline recruitment of brain regions implicated in body ownership and agency, providing evidence that they play a role in core self representation.

However, the role of the insula is likely a bottom-up integration of multisensory and interoceptive signals, a process that might be not be operative in my paradigm. Rather than expressing skepticism of the insula's importance for core self, I suggest that its role is more effectively elucidated by psychophysical paradigms like the RHI in which body-representation is disrupted from the bottom-up (c.f., Tsakiris, 2010, for a discussion on bottom-up vs. top-down processing in the RHI).

Northoff, Qin, & Feinberg (2011) expound upon this issue in terms of process-based versus content-based studies of the self. Process-based approaches, such as the one used in the current experiment, use some cognitive judgment to invoke self representations. One drawback to process-based approaches is that the cognitive task is inevitably confounded with representational content. My use of the semantic task ameliorated this issue to an extent, but the need for multiple complementary approaches to the study of person representations is paramount.

Future Directions

The two experiments I have carried out on the subject of core and conceptual representation provide a foundation upon which future studies can build. My findings raise many new questions; I will raise two of them here, and suggest possible approaches towards answering them.

How do body ownership, agency, and the first-person perspective contribute to an overall sense of self? The current results demonstrate that brain networks involved in body ownership, agency, and first-person perspective contribute to core self, but the nature of that contribution remains to be seen. Methods that disrupt one or more facets of core self have associated those facets with particular brain networks (e.g., Tsakiris, 2010); now, they should be performed with an eye towards a bigger picture of core and conceptual self representation. For example, how does the rubber hand illusion affect the self-reference effect in memory? How might disrupting neural signaling in TPJ or SMA with transcranial magnetic stimulation (TMS) or cathodal transcranial direct current stimulation (tDCS) affect personality measures, or the retrieval of autobiographical memories? Can the aversive nature of threatening stimuli be mollified by manipulating core self? Gallagher (2013) described a multifaceted pattern theory of self, in which selfhood emerges from the complex interactions of core and conceptual components. Empirical studies demonstrating these interactions, by grounding higher order cognitive aspects of the self in the embodied processes underlying core self is an exciting potential avenue of study.

How do core and conceptual representations differ for in-groups and out-groups? Adults and children provided a useful first step towards understanding core and conceptual representations for groups of people, but the field of group representation has barely begun to be explored. One fruitful avenue will be to explore the in-group out-group distinction. As mentioned earlier, there is emerging evidence that groups and individuals might share representational mechanisms beyond the minimal overlap I have found (Contreras et al., 2013), and that the degree of overlap may be related to the degree to which one identifies with the group in question (Scheepers et al., 2013). Having participants rate their inclusion in a group, similar to my using the IOS to measure closeness to others, is one option. However, in-group membership is notoriously susceptible to ad-hoc manipulation, providing opportunities for carefully controlled experimentation (Ellemers, 2012).

Conclusions

I have argued that core self-representation is the result of a coordinated effort among the neurocognitive systems that represent body ownership, agency, the first-person perspective, and basic sensorimotor features. Likewise, there is no single module for conceptual self representation: conceptual self emerges from a network of neural systems that, outside the context of person representation, subserve many other functions such as sensory perception and imagery, action planning, emotion, memory, decision making, and attention.

To an extent that the information they provide is available and relevant, self-representations serve as a template for representing other persons. As a consequence there is significant overlap, but also revealing differences. Representations for groups of people are less robust than for individuals: they rely heavily on visual characteristics and conceptual information drawn from semantic memory. By contrast, conceptual representations for individuals rely on person-specific information processed in cortical midline structures and the anterior temporal lobes.

The current experiment presents a broad framework for understanding person representation in the brain. It is my hope that future research will build upon these ideas, honing them and refining them, towards developing a detailed understanding of the social mind and brain.

References

Abu-Akel, A., & Shamay-Tsoory, S. (2011). Neuroanatomical and neurochemical bases of theory of mind. *Neuropsychologia*, *49*(11), 2971–2984. http://doi.org/10.1016/j.neuropsychologia.2011.07.012

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277. http://doi.org/10.1038/nrn1884

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences*, *33*(04), 245–266. http://doi.org/10.1017/S0140525X10000853

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*(4), 596.

Arzy, S., Thut, G., Mohr, C., Michel, C. M., & Blanke, O. (2006). Neural Basis of Embodiment: Distinct Contributions of Temporoparietal Junction and Extrastriate Body Area. *Journal of Neuroscience*, *26*(31), 8074–8081. http://doi.org/10.1523/JNEUROSCI.0745-06.2006

Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nature Neuroscience*, *7*(5), 542–548. http://doi.org/10.1038/nn1241

Aziz-Zadeh, L., Sheng, T., Liew, S.-L., & Damasio, H. (2012). Understanding Otherness: The Neural Bases of Action Comprehension and Pain Empathy in a Congenital Amputee. *Cerebral Cortex*, *22*(4), 811–819. http://doi.org/10.1093/cercor/bhr139

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

Barsalou, L.W. (1983). Ad hoc categories. *Memory & Cognition, 11,* 211-227.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(4), 629–654.

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, *59*(1), 617–645. http://doi.org/10.1146/annurev.psych.59.103006.093639

Barsalou, L. W. (2013). Ad hoc categories. *Memory & Cognition*, *11*(3), 211–227.

Baumeister, R. F. (2010). The self. In *Advanced social psychology: The state of science* (pp. 680–740).

Beckes, L., Coan, J. A., & Hasselmo, K. (2013). Familiarity promotes the blurring of self and other in the neural representation of threat. *Social Cognitive and Affective Neuroscience*, *8*(6), 670–677. http://doi.org/10.1093/scan/nss046

Blakemore, S., & Frith, C. (2003). Self-awareness and action. *Current Opinion in Neurobiology*, *13*(2), 219–224. http://doi.org/10.1016/S0959-4388(03)00043-6

Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience*. http://doi.org/10.1038/nrn3292

Blanke, O., & Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Sciences*, *13*(1), 7–13. http://doi.org/10.1016/j.tics.2008.10.003

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The Brain's Default Network: Anatomy, Function, and Relevance to Disease. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. http://doi.org/10.1196/annals.1440.011

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*(2), 49–57. http://doi.org/10.1016/j.tics.2006.11.004

Carey, S., & Xu, F. (2001). Infants’ knowledge of objects: Beyond object files and object tracking. *Cognition*, *80*(1), 179–213.

Carluer, L., Mondou, A., Buhour, M.-S., Laisney, M., Pélerin, A., Eustache, F., … Desgranges, B. (2015). Neural substrate of cognitive theory of mind impairment in amyotrophic lateral sclerosis. *Cortex*, *65*, 19–30. http://doi.org/10.1016/j.cortex.2014.12.010

Catrambone, R., Beike, D., & Niedenthal, P. (1996). Is the Self-Concept a Habitual Referent in Judgments of Similarity? *Psychological Science*, *7*(3), 158–163. http://doi.org/10.1111/j.1467-9280.1996.tb00349.x

Cavanna, A. E. & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, *129*(3), 564–583. http://doi.org/10.1093/brain/awl004

Christoff, K., Cosmelli, D., Legrand, D., & Thompson, E. (2011). Specifying the self for cognitive neuroscience. *Trends in Cognitive Sciences*, *15*(3), 104–112. http://doi.org/10.1016/j.tics.2011.01.001

Colzato, L. S., van den Wildenberg, W. P. M., & Hommel, B. (2013). Increasing self–other integration through divergent thinking. *Psychonomic Bulletin & Review*, *20*(5), 1011–1016. http://doi.org/10.3758/s13423-013-0413-4

Contreras, J. M., Schirmer, J., Banaji, M. R., & Mitchell, J. P. (2013). Common Brain Regions with Distinct Patterns of Neural Responses during Mentalizing about Groups and Individuals. *Journal of Cognitive Neuroscience*, *25*(9), 1406–1417. http://doi.org/10.1162/jocn_a_00403

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, *3*(3), 292–297.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *351*(1346), 1413–1420.

Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. Heinemann: London.

Damasio, A. R. (2010). *Self comes to mind: Constructing the conscious brain*. Vintage.

David, N., Cohen, M. X., Newen, A., Bewernick, B. H., Shah, N. J., Fink, G. R., & Vogeley, K. (2007). The extrastriate cortex distinguishes between the consequences of one's own and others' behavior. *NeuroImage*, *36*(3), 1004–1014. http://doi.org/10.1016/j.neuroimage.2007.03.030

De Brigard, F., Nathan Spreng, R., Mitchell, J. P., & Schacter, D. L. (2015). Neural activity associated with self, other, and object-based counterfactual thinking. *NeuroImage*. http://doi.org/10.1016/j.neuroimage.2014.12.075

Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: a social cognitive neuroscience view. *Trends in Cognitive Sciences*, *7*(12), 527–533. http://doi.org/10.1016/j.tics.2003.10.004

Dennett, D. (1993). *Consciousness explained*. Penguin UK.

Denny, B. T., Kober, H., Wager, T. D., & Ochsner, K. N. (2012). A meta-analysis of functional neuroimaging studies of self-and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *Journal of Cognitive Neuroscience, 24*(8), 1742–1752.

Dobbins, I. G., & Wagner, A. D. (2005). Domain-general and Domain-sensitive Prefrontal Mechanisms for Recollecting Events and Detecting Novelty. *Cerebral Cortex*, *15*(11), 1768–1778. http://doi.org/10.1093/cercor/bhi054

Dosenbach, N. U. F., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., … Petersen, S. E. (2006). A Core System for the Implementation of Task Sets. *Neuron*, *50*(5), 799–812. http://doi.org/10.1016/j.neuron.2006.04.031

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.

Drane, D. L., Ojemann, G. A., Aylward, E., Ojemann, J. G., Johnson, L. C., Silbergeld, D. L., … Tranel, D. (2008). Category-specific naming and recognition deficits in temporal lobe epilepsy surgical patients. *Neuropsychologia*, *46*(5), 1242–1255.

Drucker, J. H., Wilson-Mendenhall, C. D., Barrett, L. F., & Barsalou, L. W. (submitted). Establishing the neural bases of core self and conceptual self.

Ehrsson, H. H. (2005). Touching a Rubber Hand: Feeling of Body Ownership Is Associated with Activity in Multisensory Brain Areas. *Journal of Neuroscience*, *25*(45), 10564–10573. http://doi.org/10.1523/JNEUROSCI.0800-05.2005

Ehrsson, H. H., Spence, C., & Passingham, R. E. (2004). That’s my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science*, *305*(5685), 875–877.

Ellemers, N. (2012). The Group Self. *Science*, *336*(6083), 848–852. http://doi.org/10.1126/science.1220987

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, *12*(10), 388–396. http://doi.org/10.1016/j.tics.2008.07.004

Esslen, M., Metzler, S., Pascual-Marqui, R., & Jancke, L. (2008). Pre-reflective and reflective self-reference: A spatiotemporal EEG analysis. *NeuroImage*, *42*(1), 437–449. http://doi.org/10.1016/j.neuroimage.2008.01.060

Farb, N. A. S., Segal, Z. V., Mayberg, H., Bean, J., McKeon, D., Fatima, Z., & Anderson, A. K. (2007). Attending to the present: mindfulness meditation reveals distinct neural modes of self-reference. *Social Cognitive and Affective Neuroscience*, *2*(4), 313–322. http://doi.org/10.1093/scan/nsm030

Farrer, C., Frey, S. H., Horn, J. D. V., Tunik, E., Turk, D., Inati, S., & Grafton, S. T. (2008). The Angular Gyrus Computes Action Awareness Representations. *Cerebral Cortex*, *18*(2), 254–261. http://doi.org/10.1093/cercor/bhm050

Fotopoulou, A., Pernigo, S., Maeda, R., Rudd, A., & Kopelman, M. A. (2010). Implicit awareness in anosognosia for hemiplegia: unconscious interference without conscious re-representation. *Brain*, *133*(12), 3564–3577. http://doi.org/10.1093/brain/awq233

Freton, M., Lemogne, C., Bergouignan, L., Delaveau, P., Lehéricy, S., & Fossati, P. (2014). The eye of the self: precuneus volume and visual perspective during autobiographical memory retrieval. *Brain Structure and Function*, *219*(3), 959–968. http://doi.org/10.1007/s00429-013-0546-2

Fuster, J. M. (2001). The prefrontal cortex—an update: time is of the essence. *Neuron*, *30*(2), 319–333.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences*, *4*(1), 14–21.

Gallagher, S. (2013). A Pattern Theory of Self. *Frontiers in Human Neuroscience*, *7*. http://doi.org/10.3389/fnhum.2013.00443

Gao, T., & Scholl, B. J. (2010). Are objects required for object-files? Roles of segmentation and spatiotemporal continuity in computing object persistence. *Visual Cognition*, *18*(1), 82–109. http://doi.org/10.1080/13506280802614966

Gawryluk, J. R., Mazerolle, E. L., & D'Arcy, R. C. N. (2014). Does functional MRI detect activation in white matter? A review of emerging evidence, issues, and future directions. *Brain Imaging Methods*, *8*, 239. http://doi.org/10.3389/fnins.2014.00239

Gecas, V. (1982). The self-concept. *Annual Review of Sociology*, *8*, 1–33.

Gillihan, S. J., & Farah, M. J. (2005). Is Self Special? A Critical Review of Evidence From Experimental Psychology and Cognitive Neuroscience. *Psychological Bulletin*, *131*(1), 76–97. http://doi.org/10.1037/0033-2909.131.1.76

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25. http://doi.org/10.1016/0166-2236(92)90344-8

Gordon, R. D., & Irwin, D. E. (1996). What's in an object file? Evidence from priming studies. *Perception & Psychophysics*, *58*(8), 1260–1277.

Gross, R. G., & Grossman, M. (2008). Update on apraxia. *Current Neurology and Neuroscience Reports*, *8*(6), 490–496. http://doi.org/10.1007/s11910-008-0078-y

Gruberger, M., Levkovitz, Y., Hendler, T., Harel, E. V., Harari, H., Ben Simon, E., … Zangen, A. (2015). I think therefore I am: Rest-related prefrontal cortex neural activity is involved in generating the sense of self. *Consciousness and Cognition*, *33*, 414–421. http://doi.org/10.1016/j.concog.2015.02.008

Haggard, P., & Chambon, V. (2012). Sense of agency. *Current Biology*, *22*(10), R390–R392.

Haggard, P., & Whitford, B. (2004). Supplementary motor area provides an efferent signal for sensory suppression. *Cognitive Brain Research*, *19*(1), 52–58. http://doi.org/10.1016/j.cogbrainres.2003.10.018

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

Harter, S. (1983). Development perspectives on the self-system. *Handbook of child psychology: formerly Carmichael's Manual of child psychology/Paul H. Mussen, editor*.

Heatherton, T. F. (2011). Neuroscience of Self and Self-Regulation. *Annual Review of Psychology*, *62*(1), 363–390. http://doi.org/10.1146/annurev.psych.121208.131616

Heydrich, L., & Blanke, O. (2013). Distinct illusory own-body perceptions caused by damage to posterior insula and extrastriate cortex. *Brain*, *136*(3), 790–803. http://doi.org/10.1093/brain/aws364

Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience & Biobehavioral Reviews*, *34*(4), 575–583. http://doi.org/10.1016/j.neubiorev.2009.11.007

Hickok, G. (2013). Do Mirror Neurons Subserve Action Understanding? *Neuroscience Letters*, *540*, 56–58. http://doi.org/10.1016/j.neulet.2012.11.001

Hoffman, P., Evans, G. A. L., & Lambon Ralph, M. A. (2014). The anterior temporal lobes are critically involved in acquiring new conceptual knowledge: Evidence for impaired feature integration in semantic dementia. *Cortex*, *50*, 19–31. http://doi.org/10.1016/j.cortex.2013.10.006

Hommel, B. (2004). Event files: feature binding in and across perception and action. *Trends in Cognitive Sciences*, *8*(11), 494–500. http://doi.org/10.1016/j.tics.2004.08.007

Isoda, M., & Noritake, A. (2013). What makes the dorsomedial frontal cortex active during reading the mental states of others? *Frontiers in Neuroscience*, *7*. http://doi.org/10.3389/fnins.2013.00232

James, W. (1892). *Psychology*. New York: Henry Holt & Company.

Jeannerod, M. (2004). Visual and action cues contribute to the self-other distinction. *Nature Neuroscience*, *7*(5), 422–423.

Jenkins, A. C., & Mitchell, J. P. (2011). Medial prefrontal cortex subserves diverse forms of self-reflection. *Social Neuroscience*, *6*(3), 211–218. http://doi.org/10.1080/17470919.2010.507948

Johnson, M., & Munakata, Y. (2005). Processes of change in brain and cognitive development. *Trends in Cognitive Sciences*, *9*(3), 152–158. http://doi.org/10.1016/j.tics.2005.01.009

Jordan, K. E., Clark, K., & Mitroff, S. R. (2010). See an object, hear an object file: Object correspondence transcends sensory modality. *Visual Cognition*, *18*(4), 492–503. http://doi.org/10.1080/13506280903338911

Kang, P., Lee, J., Sul, S., & Kim, H. (2013). Dorsomedial prefrontal cortex activity predicts the accuracy in estimating others' preferences. *Frontiers in Human Neuroscience*, *7,* 686. http://doi.org/10.3389/fnhum.2013.00686

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–4311.

Karnath, H.-O., & Baier, B. (2010). Right insula for our sense of limb ownership and self-awareness of actions. *Brain Structure and Function*, *214*(5-6), 411–417. http://doi.org/10.1007/s00429-010-0250-4

Kelley, W. M., Macrae, C. N., Wyland, C. N., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*(5), 785–794.

Keysers, C., & Gazzola, V. (2010). Social Neuroscience: Mirror Neurons Recorded in Humans. *Current Biology*, *20*(8), R353–R354. http://doi.org/10.1016/j.cub.2010.03.013

Klein, S. B., & Gangi, C. E. (2010). The multiplicity of self: neuropsychological evidence and its implications for the self as a construct in psychological research. *Annals of the New York Academy of Sciences*, *1191*(1), 1–15. http://doi.org/10.1111/j.1749-6632.2010.05441.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*(9), 635–642.

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural systems shared by visual imagery and visual perception: A positron emission tomography study. *NeuroImage*, *6*(4), 320–334.

Krienen, F. M., Tu, P., & Buckner, R. L. (2010). Clan Mentality: Evidence That the Medial Prefrontal Cortex Responds to Close Others. *The Journal of Neuroscience*, *30*(41), 13906–13915. http://doi.org/10.1523/JNEUROSCI.2180-10.2010

Land, M. F. (2014). Do we have an internal model of the outside world? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1636), 20130045–20130045. http://doi.org/10.1098/rstb.2013.0045

Legrand, D. (2006). The Bodily Self: The Sensori-Motor Roots of Pre-Reflective Self-Consciousness. *Phenomenology and the Cognitive Sciences*, *5*(1), 89–118. http://doi.org/10.1007/s11097-005-9015-6

Legrand, D., & Ruby, P. (2009). What is self-specific? Theoretical investigation and critical review of neuroimaging results. *Psychological Review*, *116*(1), 252–282. http://doi.org/10.1037/a0014172

Li, P., Ogura, T., Barner, D., Yang, S.-J., & Carey, S. (2009). Does the conceptual distinction between singular and plural sets depend on language? *Developmental Psychology*, *45*(6), 1644–1653. http://doi.org/10.1037/a0015553

Limanowski, J., & Blankenburg, F. (2015). Network activity underlying the illusory self-attribution of a dummy arm: Network Activity of Illusory Self-Attribution. *Human Brain Mapping*, n/a–n/a. http://doi.org/10.1002/hbm.22770

Lingnau, A., & Downing, P. E. (2015). The lateral occipitotemporal cortex in action. *Trends in Cognitive Sciences, in press*. http://doi.org/10.1016/j.tics.2015.03.006

Lombardo, M. V., & Baron-Cohen, S. (2011). The role of the self in mindblindness in autism. *Consciousness and Cognition*, *20*(1), 130–140. http://doi.org/10.1016/j.concog.2010.09.006

Lombardo, M. V., Chakrabarti, B., Bullmore, E. T., Wheelwright, S. J., Sadek, S. A., Suckling, J., … others. (2010). Shared neural circuits for mentalizing about the self and others. *Journal of Cognitive Neuroscience*, *22*(7), 1623–1635.

Ma, N., Vandekerckhove, M., Van Hoeck, N., & van Overwalle, F. (2012). Distinct recruitment of temporo-parietal junction and medial prefrontal cortex in behavior understanding and trait identification. *Social Neuroscience*, *7*(6), 591–605. http://doi.org/10.1080/17470919.2012.686925

Maddock, R. J., Garrett, A. S., & Buonocore, M. H. (2001). Remembering familiar people: the posterior cingulate cortex and autobiographical memory retrieval. *Neuroscience*, *104*(3), 667–676.

Mahayana, I. T., Tcheang, L., Chen, C.-Y., Juan, C.-H., & Muggleton, N. G. (2014). The Precuneus and Visuospatial Attention in Near and far Space: A Transcranial Magnetic Stimulation Study. *Brain Stimulation*, *7*(5), 673–679. http://doi.org/10.1016/j.brs.2014.06.012

Markus, H., Smith, J., & Moreland, R. L. (1985). Role of the self-concept in the perception of others. *Journal of Personality and Social Psychology*, *49*(6), 1494–1512. http://doi.org/10.1037/0022-3514.49.6.1494

Markus, H., & Wurf, E. (1987). The Dynamic Self-Concept: A Social Psychological Perspective. *Annual Review of Psychology*, *38*(1), 299–337. http://doi.org/10.1146/annurev.ps.38.020187.001503

Martin, A. (2007). The Representation of Object Concepts in the Brain. *Annual Review of Psychology*, *58*(1), 25–45. http://doi.org/10.1146/annurev.psych.57.102904.190143

Mashek, D. J., Aron, A., & Boncimino, M. (2003). Confusions of Self With Close Others. *Personality and Social Psychology Bulletin*, *29*(3), 382–392. http://doi.org/10.1177/0146167202250220

Mazerolle, E. L., Beyea, S. D., Gawryluk, J. R., Brewer, K. D., Bowen, C. V., & D'Arcy, R. C. N. (2010). Confirming white matter fMRI activation in the corpus callosum: Co-localization with DTI tractography. *NeuroImage*, *50*(2), 616–621. http://doi.org/10.1016/j.neuroimage.2009.12.102

McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory & Cognition*, *6*(4), 462–472.

McCrae, R. R., & Costa, P. T. (2003). *Personality in adulthood: A five-factor theory perspective*. Guilford Press.

Miller, B. L., Seeley, W. W., Mychack, P., Rosen, H. J., Mena, I., & Boone, K. (2001). Neuroanatomy of the self Evidence from patients with frontotemporal dementia. *Neurology*, *57*(5), 817–821.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202.

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research*, *6*(1), 57–77. http://doi.org/10.1016/0166-4328(82)90081-X

Mitchell, J. P. (2008). Activity in Right Temporo-Parietal Junction is Not Selective for Theory-of-Mind. *Cerebral Cortex*, *18*(2), 262–271. http://doi.org/10.1093/cercor/bhm051

Mitchell, J. P. (2009). Inferences about mental states. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1309–1316. http://doi.org/10.1098/rstb.2008.0318

Mitchell, J. P. (2009). Social psychology as a natural kind. *Trends in Cognitive Sciences*, *13*(6), 246–251. http://doi.org/10.1016/j.tics.2009.03.008

Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences*, *99*(23), 15238–15243. http://doi.org/10.1073/pnas.232395699

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable Medial Prefrontal Contributions to Judgments of Similar and Dissimilar Others. *Neuron*, *50*(4), 655–663. http://doi.org/10.1016/j.neuron.2006.03.040

Moore, J. W., Ruge, D., Wenke, D., Rothwell, J., & Haggard, P. (2010). Disrupting the experience of control in the human brain: pre-supplementary motor area contributes to the sense of agency. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1693), 2503–2509. http://doi.org/10.1098/rspb.2010.0404

Moran, J. M., Kelley, W. M., & Heatherton, T. F. (2013). What Can the Organization of the Brain's Default Mode Network Tell us About Self-Knowledge? *Frontiers in Human Neuroscience, 7*. http://doi.org/10.3389/fnhum.2013.00391

Moran, J. M., Lee, S. M., & Gabrieli, J. D. E. (2010). Dissociable Neural Systems Supporting Knowledge about Human Character and Appearance in Ourselves and Others. *Journal of Cognitive Neuroscience*, *23*(9), 2222–2230. http://doi.org/10.1162/jocn.2010.21580

Murray, R. J., Schaer, M., & Debbané, M. (2012). Degrees of separation: A quantitative neuroimaging meta-analysis investigating self-specificity and shared neural activation between self- and other-reflection. *Neuroscience & Biobehavioral Reviews*, *36*(3), 1043–1059. http://doi.org/10.1016/j.neubiorev.2011.12.013

Myers, A., & Sowden, P. T. (2008). Your hand or mine? The extrastriate body area. *NeuroImage*, *42*(4), 1669–1677. http://doi.org/10.1016/j.neuroimage.2008.05.045

Nahab, F. B., Kundu, P., Gallea, C., Kakareka, J., Pursley, R., Pohida, T., … Hallett, M. (2011). The Neural Processes Underlying Self-Agency. *Cerebral Cortex*, *21*(1), 48–55. http://doi.org/10.1093/cercor/bhq059

Nassi, J. J., & Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nature Reviews Neuroscience*, *10*(5), 360–372. http://doi.org/10.1038/nrn2619

Neisser, U. (1988). Five kinds of self-knowledge. *Philosophical Psychology*, *1*(1), 35–59. http://doi.org/10.1080/09515088808572924

Noles, N. S., Scholl, B. J., & Mitroff, S. R. (2005). The persistence of object file representations. *Perception & Psychophysics*, *67*(2), 324–334.

Northoff, G., & Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences*, *8*(3), 102–107. http://doi.org/10.1016/j.tics.2004.01.004

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *NeuroImage*, *31*(1), 440–457. http://doi.org/10.1016/j.neuroimage.2005.12.002

Northoff, G., Qin, P., & Feinberg, T. E. (2011). Brain imaging of the self – Conceptual, anatomical and methodological issues. *Consciousness and Cognition*, *20*(1), 52–63. http://doi.org/10.1016/j.concog.2010.09.011

Northoff, G., Schneider, F., Rotte, M., Matthiae, C., Tempelmann, C., Wiebking, C., … Panksepp, J. (2009). Differential parametric modulation of self-relatedness and emotions in different brain regions. *Human Brain Mapping*, *30*(2), 369–382. http://doi.org/10.1002/hbm.20510

Ollinger, J. M., Corbetta, M., & Shulman, G. L. (2001a). Separating Processes within a Trial in Event-Related Functional MRI. *NeuroImage*, *13*(1), 218–229. http://doi.org/10.1006/nimg.2000.0711

Ollinger, J. M., Corbetta, M., & Shulman, G. L. (2001b). Separating Processes within a Trial in Event-Related Functional MRI: II. Analysis. *NeuroImage*, *13*(1), 218–229. http://doi.org/10.1006/nimg.2000.0711

Ollinger, J. M., Shulman, G. L., & Corbetta, M. (2001a). Separating Processes within a Trial in Event-Related Functional MRI. *NeuroImage*, *13*(1), 210–217. http://doi.org/10.1006/nimg.2000.0710

Ollinger, J. M., Shulman, G. L., & Corbetta, M. (2001b). Separating Processes within a Trial in Event-Related Functional MRI: I. The Method. *NeuroImage*, *13*(1), 210–217. http://doi.org/10.1006/nimg.2000.0710

Perry, C. J., & Fallah, M. (2014). Feature integration and object representations along the dorsal stream visual hierarchy. *Frontiers in Computational Neuroscience, 8*, 84. http://doi.org/10.3389/fncom.2014.00084

Petrini, K., Piwek, L., Crabbe, F., Pollick, F. E., & Garrod, S. (2014). Look at those two!: The precuneus role in unattended third-person perspective of social interactions: Precuneus and Third-Person Interaction. *Human Brain Mapping*, *35*(10), 5190–5203. http://doi.org/10.1002/hbm.22543

Pollen, D. A. (2011). On the Emergence of Primary Visual Perception. *Cerebral Cortex*, *21*(9), 1941–1953. http://doi.org/10.1093/cercor/bhq285

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38. http://doi.org/10.1016/j.neuroscience.2005.06.005

Powell, L. J., Macrae, C. N., Cloutier, J., Metcalfe, J., & Mitchell, J. P. (2010). Dissociable neural substrates for agentic versus conceptual representations of self. *Journal of Cognitive Neuroscience*, *22*(10), 2186–2197.

Qin, P., & Northoff, G. (2011). How is our self related to midline regions and the default-mode network? *NeuroImage*, *57*(3), 1221–1233. http://doi.org/10.1016/j.neuroimage.2011.05.028

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.

Robbins, P. (2002). The Paradox of Self–Consciousness Revisited. *Pacific Philosophical Quarterly*, *83*(4), 424–443.

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*(9), 677.

Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nature Neuroscience*, *4*(5), 546–550. http://doi.org/10.1038/87510

Rupert, R. D. (2011). Embodiment, Consciousness, and the Massively Representational Mind. *Philosophical Topics*, *39*(1), 99–120.

Sandberg, C., Sebastian, R., & Kiran, S. (2012). Typicality mediates performance during category verification in both ad-hoc and well-defined categories. *Journal of Communication Disorders*, *45*(2), 69–83. http://doi.org/10.1016/j.jcomdis.2011.12.004

Sathian, K., Buxbaum, L. J., Cohen, L. G., Krakauer, J. W., Lang, C. E., Corbetta, M., & Fitzpatrick, S. M. (2011). Neurological Principles and Rehabilitation of Action Disorders Common Clinical Deficits. *Neurorehabilitation and Neural Repair*, *25*(5 suppl), 21S–32S. http://doi.org/10.1177/1545968311410941

Saxe, R., Jamal, N., & Powell, L. (2005). My Body or Yours? The Effect of Visual Perspective on Cortical Body Representations. *Cerebral Cortex*, *16*(2), 178–182. http://doi.org/10.1093/cercor/bhi095

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842. http://doi.org/10.1016/S1053-8119(03)00230-1

Saxe, R., Moran, J. M., Scholz, J., & Gabrieli, J. (2006). Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. *Social Cognitive and Affective Neuroscience*, *1*(3), 229–234. http://doi.org/10.1093/scan/nsl034

Saxe, R., Xiao, D.-K., Kovacs, G., Perrett, D. ., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, *42*(11), 1435–1446. http://doi.org/10.1016/j.neuropsychologia.2004.04.015

Scheepers, D., Derks, B., Nieuwenhuis, S., Lelieveld, G.-J., Van Nunspeet, F., Rombouts, S. A. R. B., & de Rover, M. (2013). The neural correlates of in-group and self-face perception: is there overlap for high identifiers? *Frontiers in Human Neuroscience*, *7*. http://doi.org/10.3389/fnhum.2013.00528

Schneider, F., Bermpohl, F., Heinzel, A., Rotte, M., Walter, M., Tempelmann, C., … Northoff, G. (2008). The resting brain and our self: Self-relatedness modulates resting state neural activity in cortical midline structures. *Neuroscience*, *157*(1), 120–131. http://doi.org/10.1016/j.neuroscience.2008.08.014

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1), 1–46.

Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind & Language*, *22*(5), 563–591.

Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychological Science*, *16*(2), 114–122.

Shen, D., & Mondor, T. A. (2008). Object file continuity and the auditory attentional blink. *Perception & Psychophysics*, *70*(5), 896–915. http://doi.org/10.3758/PP.70.5.896

Sierra, M., & David, A. S. (2011). Depersonalization: A selective impairment of self-awareness. *Consciousness and Cognition*, *20*(1), 99–108. http://doi.org/10.1016/j.concog.2010.10.018

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. http://doi.org/10.1016/j.neuropsychologia.2007.05.002

Simmons, W. K., Reddish, M., Bellgowan, P. S. F., & Martin, A. (2010). The Selectivity and Functional Connectivity of the Anterior Temporal Lobes. *Cerebral Cortex*, *20*(4), 813–825. http://doi.org/10.1093/cercor/bhp149

Spencer, M. D., Holt, R. J., Chura, L. R., Suckling, J., Calder, A. J., Bullmore, E. T., & Baron-Cohen, S. (2011). A novel functional brain imaging endophenotype of autism: the neural response to facial expression of emotion. *Translational Psychiatry*, *1*(7), e19. http://doi.org/10.1038/tp.2011.18

Stevens, W. D., Tessler, M. H., Peng, C. S., & Martin, A. (2015). Functional connectivity constrains the category-related organization of human ventral occipitotemporal cortex:

Connectivity Constrains VOTC Organization. *Human Brain Mapping*, n/a–n/a. http://doi.org/10.1002/hbm.22764

Sugiura, M., Shah, N., Zilles, K., & Fink, G. (2005). Cortical representations of personally familiar objects and places: functional organization of the human posterior cingulate cortex. *Cognitive Neuroscience, Journal of*, *17*(2), 183–198.

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia*, *44*(12), 2189–2208. http://doi.org/10.1016/j.neuropsychologia.2006.05.023

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: a meta-analysis. *Psychological Bulletin*, *121*(3), 371.

Tagini, A., & Raffone, A. (2010). The “I” and the “Me” in self-referential awareness: a neurocognitive hypothesis. *Cognitive Processing*, *11*(1), 9–20. http://doi.org/10.1007/s10339-009-0336-1

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring “how” from “where.” *Neuropsychologia*, *41*(3), 280–292.

Tsakiris, M. (2010). My body in the brain: A neurocognitive model of body-ownership. *Neuropsychologia*, *48*(3), 703–712. http://doi.org/10.1016/j.neuropsychologia.2009.09.034

Tsakiris, M., Hesse, M. D., Boy, C., Haggard, P., & Fink, G. R. (2007). Neural Signatures of Body Ownership: A Sensory Network for Bodily Self-Consciousness. *Cerebral Cortex*, *17*(10), 2235–2244. http://doi.org/10.1093/cercor/bhl131

Tsakiris, M., Longo, M. R., & Haggard, P. (2010). Having a body versus moving your body: Neural signatures of agency and body-ownership. *Neuropsychologia*, *48*(9), 2740–2749. http://doi.org/10.1016/j.neuropsychologia.2010.05.021

Tsakiris, M., Schütz-Bosbach, S., & Gallagher, S. (2007). On agency and body-ownership: Phenomenological and neurocognitive reflections. *Consciousness and Cognition*, *16*(3), 645–660. http://doi.org/10.1016/j.concog.2007.05.012

Uddin, L. Q., Iacoboni, M., Lange, C., & Keenan, J. P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends in Cognitive Sciences*, *11*(4), 153–157. http://doi.org/10.1016/j.tics.2007.01.001

Urgen, B. A., & Miller, L. E. (2015). Towards an Empirically Grounded Predictive Coding Account of Action Understanding. *Journal of Neuroscience*, *35*(12), 4789–4791. http://doi.org/10.1523/JNEUROSCI.0144-15.2015

Vallar, G., & Ronchi, R. (2009). Somatoparaphrenia: a body delusion. A review of the neuropsychological literature. *Experimental Brain Research*, *192*(3), 533–551. http://doi.org/10.1007/s00221-008-1562-y

Van der Meer, L., Costafreda, S., Aleman, A., & David, A. S. (2010). Self-reflection and the brain: A theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia. *Neuroscience & Biobehavioral Reviews*, *34*(6), 935–946. http://doi.org/10.1016/j.neubiorev.2009.12.004

Vanderwal, T., Hunyadi, E., Grupe, D. W., Connors, C. M., & Schultz, R. T. (2008). Self, mother and abstract other: An fMRI study of reflective social processing. *NeuroImage*, *41*(4), 1437–1446. http://doi.org/10.1016/j.neuroimage.2008.03.058

Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience, 10*(11), 792–802. http://doi.org/10.1038/nrn2733

van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, *30*(3), 829–858. http://doi.org/10.1002/hbm.20547

van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage*, *48*(3), 564–584. http://doi.org/10.1016/j.neuroimage.2009.06.009

Vogeley, K., & Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences*, *7*(1), 38–42. http://doi.org/10.1016/S1364-6613(02)00003-7

Von Arx, S. W., Müri, R. M., Heinemann, D., Hess, C. W., & Nyffeler, T. (2010). Anosognosia for cerebral achromatopsia—A longitudinal case study. *Neuropsychologia*, *48*(4), 970–977. http://doi.org/10.1016/j.neuropsychologia.2009.11.018

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron*, *31*(2), 329–338.

Wagner, D. D., Haxby, J. V., & Heatherton, T. F. (2012). The representation of self and person knowledge in the medial prefrontal cortex. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(4), 451–470. http://doi.org/10.1002/wcs.1183

Weiner, K. S., & Grill-Spector, K. (2011). Not one extrastriate body area: Using anatomical landmarks, hMT+, and visual field maps to parcellate limb-selective activations in human lateral occipitotemporal cortex. *NeuroImage*, *56*(4), 2183–2199. http://doi.org/10.1016/j.neuroimage.2011.03.041

Whitfield-Gabrieli, S., Moran, J. M., Nieto-Castañón, A., Triantafyllou, C., Saxe, R., & Gabrieli, J. D. E. (2011). Associations and dissociations between default and self-reference networks in the human brain. *NeuroImage*, *55*(1), 225–232. http://doi.org/10.1016/j.neuroimage.2010.11.048

Woodman, G. F., & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review*, *11*(2), 269–274.

Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, *13*(4), 167–174. http://doi.org/10.1016/j.tics.2009.01.008

Zmigrod, S., & Hommel, B. (2009). Auditory event files: Integrating auditory perception and action planning. *Attention, Perception & Psychophysics*, *71*(2), 352–362. http://doi.org/10.3758/APP.71.2.352

Tables

Table 1. Results from the core representation conjunction across individuals in Phase 1 (I, Relative, Friend, and Acquaintance), showing activations shared across all four individuals, shared across subsets of three individuals, shared across subsets of two individuals, and unique for each individual.

		All 4 individuals			
Region	**BA**	**Volume**	**CM**		
R extrastriate ctx.	19, 18	**108**	26	-81	4
L striate & extrastriate ctx.	18, 17	**45**	-23	-84	10
R fusiform gyrus	37	**18**	41	-52	-10
R pMTG	39	**16**	40	-70	23
L extrastriate ctx.	18	**11**	-31	-73	-8
R precuneus	7	**9**	9	-56	60
R sup. par. lobule	19	**2**	35	-75	42

	I * Relative * Friend					**I * Relative * Acquaintance**				
Region	**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
R MTG	39	**25**	40	-68	20	-	-	-	-	-
L MTG	39	**19**	-34	-77	23	-	-	-	-	-
R precuneus	-	-	-	-	-	7	**11**	5	-56	54

	I * Friend * Acquaintance					**Friend * Relative * Acquaintance**				
Region	**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
R fusiform gyrus	37	**43**	36	-58	-11	-	-	-	-	-
R MTG	19, 39	**16**	37	-72	27	-	-	-	-	-
L striate & extrastriate ctx.	-	-	-	-	-	17,18,19	**393**	-18	-88	-2
R striate & extrastriate ctx.	-	-	-	-	-	17,18,19	**280**	18	-85	-3

	I * Relative					I * Friend					I * Acquaintance				
Region	**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
L extrastriate ctx.	19	**35**	-30	-84	28	19	**29**	-33	-76	25	19	**29**	-28	-85	17
R TPJ (supramarginal gyrus)	-	-	-	-	-	40	**46**	54	-47	40					
R extrastriate ctx	-	-	-	-	-	19	**30**	42	-75	20	19	**77**	42	-63	-8
& fusiform gyrus	-	-	-	-	-	-	-	-	-	-	37, 20	included			
R MTG & angular gyrus	-	-	-	-	-	-	-	-	-	-	39	**55**	31	-74	32
R precuneus	-	-	-	-	-	-	-	-	-	-	7	**31**	5	-64	49

	Relative * Friend					Relative * Acquaintance					Friend * Acquaintance				
Region	**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
L striate ctx.	17	**13**	-8	-98	2	-	-	-	-	-					
R extrastriate ctx.	-	-	-	-	-	18	**194**	1	-87	1	19	**38**	33	-77	-11
R PCC	-	-	-	-	-	-	-	-	-	-	31	**46**	6	-38	38
R S1/M1	-	-	-	-	-	-	-	-	-	-	3, 4	**27**	34	-26	50

	I					Relative				
Region	**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
R TPJ (inf. parietal lobule),	40	**533**	41	-64	9	-	-	-	-	-
R pSTG, pSTS, pMTG,	39, 22					-	-	-	-	-
R extrastriate ctx.,	19					19	**23**	13	-94	16
R fusiform gyrus	37					-	-	-	-	-
L extrastriate, fusiform	19, 37	**379**	-37	-70	9	-	-	-	-	-
L cerebellum	NA	**54**	-17	-41	-22	-	-	-	-	-
R precuneus, sup. par. lob.	7	**46**	12	-65	56	-	-	-	-	-
L precuneus	7	**33**	-11	-54	58	-	-	-	-	-
R dlPFC	8, 9	**20**	31	22	39	8	**43**	-32	34	46
R frontopolar ctx	-	-	-	-	-	10	**98**	25	59	9

	Friend					Acquaintance				
Region	**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
R premotor ctx, dlPFC	6, 8, 9	**105**	26	20	48	-	-	-	-	-
B precuneus	7	**60**	4	-58	36	7	**43**	-5	-63	46
R fusiform &	37	**31**	33	-36	-11		*see below with MTG*			
parahippocampal gyri	36	*included*				-	-	-	-	-
R extrastriate ctx,	19	**29**	40	-74	-17	17,18,19	**445**	-8	-81	-3
L extrastriate ctx.	-	-	-	-	-		*included*			
& fusiform gyrus	-	-	-	-	-		*included*			
R sup. par. lob., precuneus	-	-	-	-	-	19, 7	**186**	24	-71	40
R S1/M1	-	-	-	-	-	3, 4	**55**	38	-25	53
R mid, posterior cingulate						31	**48**	8	-35	38
R fusiform gyrus & MTG	-	-	-	-	-	37, 20	**44**	53	-50	-13

Note. Activations in the conjunctions shown were active above the active baseline within every condition at an individual voxel threshold of $p < .005$ and a cluster threshold of 34 voxels, yielding a whole-brain threshold of $p < .05$ corrected for multiple comparisons. Note that clusters smaller than 34 voxels reflect fragments of larger supra-threshold clusters that fell into the shared overlap across conditions. Abbreviations: L: left; R: right; B: bilateral, dlPFC: dorsolateral prefrontal cortex, MTG: middle temporal gyrus, PCC: posterior cingulate cortex, S1/M1: primary somatosensory and primary motor cortices, STG: superior temporal gyrus, STS: superior temporal sulcus, TPJ: temporoparietal junction

Table 2. Significant activations in core representation pairwise contrasts between all possible pairs of individuals in Phase 1 (I, Relative, Friend, Acquaintance).

	I vs. Relative						**I vs. Friend**						**I vs. Acquaintance**					
Region	**BA**	**Volume**	**Peak t**	**(x, y, z)**			**BA**	**Volume**	**Peak t**	**(x, y, z)**			**BA**	**Volume**	**Peak t**	**(x, y, z)**		
R extrastriate ctx.,	19	**250**	7.76	40	-64	-6	19	**329**	7.93	40	-52	-6	19	**400**	8.42	38	-52	-6
R fusiform gyrus	37						37						37, 39					
L extrastriate ctx.,	19	**153**	6.00	-44	-64	0	19	**169**	5.47	-44	-62	0	19	**165**	6.44	-46	-64	0
L pMTG, pITG	37						37						37					
L pSTG	-	-	-	-	-	-	-	-	-	-	-	-	39					
L caudate & putamen	-	-	-	-	-	-	-	-	-	-	-	-	NA	**109**	4.61	-16	2	14
R TPJ (supramarginal gyrus)	40	**78**	5.15	56	-34	38	-	-	-	-	-	-	-	-	-	-	-	-
L fusiform	37	**59**	5.34	-32	-50	-10	-	-	-	-	-	-	-	-	-	-	-	-
& parahippocampal gyri	36																	
L TPJ (supramarginal gyrus),	40	**55**	4.22	-46	-32	38	-	-	-	-	-	-	-	-	-	-	-	-
L S1	3																	
L extrastriate ctx	19	**52**	5.98	-40	-82	30	-	-	-	-	-	-	19	**36**	4.19	-38	-86	18
R pMTG	39, 22	**49**	6.36	50	-46	8	-	-	-	-	-	-	-	-	-	-	-	-
R vlPFC (p. opercularis)	44	**39**	5.29	62	10	26	-	-	-	-	-	-	-	-	-	-	-	-
L mOFC	-	-	-	-	-	-	-	-	-	-	-	-	10	**69**	5.54	-32	50	2

	Relative vs. I						**Relative vs. Friend**						**Relative vs. Acquaintance**					
Region	**BA**	**Volume**	**Peak t**	**(x, y, z)**			**BA**	**Volume**	**Peak t**	**(x, y, z)**			**BA**	**Volume**	**Peak t**	**(x, y, z)**		
B striate	17,	**1455**	-10.54	-20	-92	-6	-	-	-	-	-	-	-	-	-	-	-	-
& extrastriate ctx.	18, 19						-	-	-	-	-	-	-	-	-	-	-	-

	Friend vs. I						**Friend vs. Relative**						**Friend vs. Acquaintance**					
Region	**BA**	**Volume**	**Peak t**	**(x, y, z)**			**BA**	**Volume**	**Peak t**	**(x, y, z)**			**BA**	**Volume**	**Peak t**	**(x, y, z)**		
B striate	17,	**1035**	-10.86	-16	-88	-10	-	-	-	-	-	-	-	-	-	-	-	-
& extrastriate ctx.	18, 19																	
B PCC	31	**89**	-4.72	-8	-44	30	-	-	-	-	-	-	-	-	-	-	-	-

L SMA	6	**50**	-5.77	-8	14	48	-	-	-	-	-	-	-	-	-	-	-	-
R S1	-	-	-	-	-	-	3, 2	**106**	-5.04	38	-28	48	-	-	-	-	-	-

	Acquaintance vs. I						**Acquaintance vs. Relative**						**Acquaintance vs. Friend**					
Region	**BA**	**Volume**	**Peak t**	**(x, y, z)**			**BA**	**Volume**	**Peak t**	**(x, y, z)**			**BA**	**Volume**	**Peak t**	**(x, y, z)**		
B striate	17,	**1870**	-10.16	16	-82	-10	**L** 17,18	**48**	-4.81	-10	-86	-6	17,	**672**	-8.28	-10	-86	-6
& extrastriate ctx.	18, 19						**R** 18	**79**	-6.47	14	-74	-12	18,19					
R S1	-	-	-	-	-	-	3,2	**122**	-4.69	34	-28	42						
L S1/M1	-	-	-	-	-	-	3,2,4	**111**	-5.17	-34	-26	62	-	-	-	-	-	-
L fusiform gyrus	-	-	-	-	-	-	-	-	-	-	-	-	37	**51**	-5.63	-38	-50	-24

Note. Activations in the contrasts shown were active at an individual voxel threshold of $p < .005$ and a cluster threshold of 34 voxels, yielding a whole-brain threshold of $p < .05$ corrected for multiple comparisons. Abbreviations: **L**: left; **R**: right; **B**: bilateral, **ITG**: inferior temporal gyrus, **M1**: primary motor cortex, **mOFC**: medial orbitofrontal cortex, **MTG**: middle temporal gyrus, **PCC**: posterior cingulate cortex, **S1**: primary somatosensory cortex, **SMA**: supplementary motor area, **STG**: superior temporal gyrus, **vlPFC**: ventrolateral prefrontal cortex.

Table 3. Correlations with IOS ratings for each individual (closeness to self)

BA	Region	Volume	Peak t	Peak voxel		
Individual Phase						
positive						
19, 37	**R** extrastriate ctx., fusiform gyrus	308	6.47	34	-80	12
19, 37	**L** extrastriate ctx., fusiform gyrus	109	6.40	-46	-64	0
NA	**L** caudate, putamen, & globus pallidus (ext.)	38	4.75	-16	2	14
negative						
17, 18, 19	**B** striate & extrastriate ctx.	1541	-9.78	16	-82	-10
Property Phase						
positive						
9, 24, 32	**B** dmPFC, dACC	218	5.85	-8	26	26
40	**L** TPJ (supramarginal gyrus)	97	6.54	-50	-34	26
negative						
31, 23, 30, 29	**B** PCC, **B** retrosplenial ctx.	199	-5.14	-8	-50	14
18	**L** extrastriate ctx.	53	-4.62	-14	-82	-22
36	**L** perirhinal cortex	37	-4.46	-38	-32	-10

Note. Activations in the correlations shown were active at an individual voxel threshold of $p < .005$ and a cluster threshold of 34 voxels, yielding a whole-brain threshold of $p < .05$ corrected for multiple comparisons. Abbreviations: **L**: left, **R**: right, **B**: bilateral, **dACC**: dorsal anterior cingulate cortex, **dmPFC**: dorsomedial prefrontal cortex, **PCC**: posterior cingulate cortex.

Table 4. Results from the core representation conjunction across the three tasks in Phase 1 (Individuals, Groups, and semantic Prompts), showing activations shared across all three tasks, subsets of two tasks, and unique for each task.

	Individuals * Groups * Prompts		
Region	**BA**	**Volume**	**CM**
R extrastriate ctx.	18, 19	**93**	(21, -84, -2)
L extrastriate ctx.	18, 19	**60**	(-19, -84, -3)
R fusiform g.	37	**1**	(41, -53, -13)

	Individuals * Groups			Individuals * Prompts			Groups * Prompts		
Region	**BA**	**Volume**	**CM**	**BA**	**Volume**	**CM**	**BA**	**Volume**	**CM**
R extrastriate ctx.	-	-	- - -	18, 19	**43**	(20, -79, -7)	17, 18	**27**	(20, -89, 5)
L extrastriate ctx.	-	-	- - -	17, 18	**29**	(-12, -87, -4)	17, 18, 19	**110**	(-22, -87, -2)

	Individuals			Groups			Prompts		
Region	**BA**	**Volume**	**CM**	**BA**	**Volume**	**CM**	**BA**	**Volume**	**CM**
B dmPFC, ACC, caudate	-	-	- - -	-	-	- - -	10, 32	**590**	0 39 5
R S1/M1	3, 4	**51**	30 -25 46	-	-	- - -	-	-	- - -
B precuneus	7	**66**	1 -58 42	-	-	- - -	-	-	- - -
R TPJ (supramarginal g.)	40	**35**	55 -47 40	-	-	- - -	-	-	- - -
R fusiform gyrus	37	**26**	38 -54 -9	-	-	- - -	37, 20	**77**	46 -47 -13
R parahippocampal g., STS	-	-	- - -	36, 22	**51**	34 -42 3	-	-	- - -
L striate & extrastriate ctx.	-	-	- - -	-	-	- - -	17, 18, 19	**156**	-18 -90 -6
R striate & extrastriate ctx.	-	-	- - -	-	-	- - -	17, 18	**127**	20 -89 -2
R dlPFC	-	-	- - -	-	-	- - -	8	**94**	29 28 42
L parahippocampal gyrus	-	-	- - -	-	-	- - -	36	**68**	-24 -35 4
L dlPFC	-	-	- - -	-	-	- - -	8, 9	**54**	-31 34 42
R frontopolar ctx.	-	-	- - -	-	-	- - -	10	**52**	12 63 28

Note. Activations in the conjunctions shown were active above the active baseline within every condition at an individual voxel threshold of $p < .005$ and a cluster threshold of 34 voxels, yielding a whole-brain threshold of $p < .05$ corrected for multiple comparisons. Note that clusters smaller than 34 voxels reflect fragments of larger supra-threshold clusters that fell into the shared overlap across conditions. Abbreviations: **L**: left; **R**: right; **B**:

bilateral, **ACC**: anterior cingulate cortex, **dlPFC**: dorsolateral prefrontal cortex, **dmPFC**: dorsomedial prefrontal cortex, **S1/M1**: primary somatosensory cortex and primary motor cortex, **STS**: superior temporal sulcus, **TPJ**: temporoparietal junction**.**

Table 5. Significant activations in core representation pairwise contrasts between all possible pairs of tasks in Phase 1 (Individuals, Groups, and semantic Prompts).

BA	Region	Volume	Peak t	Peak voxel		
A) Individuals > Groups						
6	**B** SMA	**190**	6.96	-8	-2	62
6	**R** premotor cortex	**153**	5.75	28	-4	50
40, 39, 7, 19	**R** TPJ (incl. inferior parietal lobule), **R** superior parietal lobule & extrastriate ctx.	**109**	5.01	52	-52	44
6	**L** premotor ctx.	**106**	5.79	-26	-2	54
40, 39, 7, 19	**L** TPJ (incl. inferior parietal lobule), **L** superior parietal lobule & extrastriate ctx.	**39**	4.15	-32	-62	42
B) Individuals > Prompts						
	none					
C) Groups > Prompts						
	none					
D) Groups > Individuals						
30	**L** retrosplenial ctx.	**56**	-4.26	-8	-50	6
E) Prompts > Individuals						
9, 10, 32, 24,	**B** dmPFC, vmPFC **B** ACC	**993**	-9.05	4	58	2
17, 18	**R** striate & extrastriate ctx.	**215**	-6.44	20	-88	-10
17, 18, 19	**L** striate & extrastriate ctx.	**206**	-5.83	-26	-86	-10
17, 18	**L** striate & extrastriate ctx.	**51**	-4.50	-4	-88	8
NA	**L** thalamus	**35**	-5.01	-14	-28	2
NA	**L** caudate head, putamen	**34**	-4.54	-10	20	6
F) Prompts > Groups						
8, 9	**R** dlPFC	**96**	-5.93	32	28	44

Note. Activations in the contrasts shown were active at an individual voxel threshold of $p < .005$ and a cluster threshold of 34 voxels, yielding a whole-brain threshold of $p < .05$ corrected for multiple comparisons. Abbreviations: **L**: left; **R**: right; **B**: bilateral, **ACC**: anterior cingulate cortex, dlPFC: dorsolateral prefrontal cortex, dmPFC: dorsomedial prefrontal cortex, **SMA**: supplementary motor area, **TPJ**: temporoparietal junction, vmPFC: ventromedial prefrontal cortex.

Table 6. Results from the conceptual representation conjunction across individuals in Phase 2 (I, Relative, Friend, and Acquaintance), showing activations shared across all four individuals, shared across subsets of three individuals, shared across subsets of two individuals, and unique for each individual.

		All 4 individuals			
Region	**BA**	**Volume**	**CM**		
B extrastriate ctx.	17, 18, 19,	**1116**	-4	-86	-4
L fusiform g.	37, 20				
L premotor ctx, dlPFC,	6, 9, 46,	**486**	-41	18	18
L vlPFC,	45, 44,				
L OFC	47, 11				
B SMA	6	**184**	-5	9	52
L dlPFC	8	**5**	-12	27	48

	I * Relative * Friend					**I * Relative * Acquaintance**				
Region	**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
L dlPFC, OFC	-	-	-	-	-	46,47,11	**39**	-45	34	5

	I * Friend * Acquaintance					**Friend * Relative * Acquaintance**				
Region	**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
L premotor ctx.,	6,	**57**	-42	4	32	-	-	-	-	-
L dlPFC,	9,									
L vlPFC (p. opercularis)	44									
B precuneus, retroplenial ctx	-	-	-	-	-	31,23	**99**	-2	-54	23
L extrastriate ctx.	-	-	-	-	-	18	**41**	-18	-97	-11

Region	I * Relative					I * Friend					I * Acquaintance				
	BA	**Volume**	**CM**			**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
R extrastriate ctx.	-	-	-	-	-	18,19	**38**	39	-81	-10	-	-	-	-	-

Region	Relative * Friend					Relative * Acquaintance					Friend * Acquaintance				
	BA	**Volume**	**CM**			**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
L STG, MTG	-	-	-	-	-	22, 21	**41**	-54	-7	-10	-	-	-	-	-
B vmPFC	-	-	-	-	-	-	-	-	-	-	10	**49**	-1	43	-6
R retrosplenial ctx., PCC		-	-	-	-	-	-	-	-	-	31,23	**38**	1	-52	24

Region	I					Relative				
	BA	**Volume**	**CM**			**BA**	**Volume**	**CM**		
L SMA, dmPFC	6, 8, 9	**130**	-11	31	48	-	-	-	-	-
R extrastriate ctx.	18	**91**	40	-84	-3	-	-	-	-	-
L MTG, fusiform g.	22,37	**46**	-48	-43	-1	-	-	-	-	-
L vlPFC, OFC	44,45,47	**44**	-50	22	4	-	-	-	-	-
L dlPFC	9,46	**40**	-53	27	25	-	-	-	-	-
R fusiform gyrus	37	**39**	36	-55	-13	-	-	-	-	-

	Friend					Acquaintance				
Region	**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
L SMA, dmPFC	6, 8, 9	**97**	0	16	39	-	-	-	-	-
B hypothalamus	NA	**50**	0	-17	-4	-	-	-	-	-
R premotor ctx.	6	**35**	25	-6	43	-	-	-	-	-
B PCC,	-	-	-	-	-	31,23,	**136**	-7	-51	25
& retrosplenial ctx.	-	-	-	-	-	30,29				
L ATL, STS	-	-	-	-	-	38,22,21	**133**	-52	-12	-7
L extrastriate ctx.	-	-	-	-	-	18,19	**92**	-2	-83	-16
L pMTG	-	-	-	-	-	39	**58**	-45	-69	25
L premotor ctx, dlPFC	-	-	-	-	-	6,8	**53**	-44	5	42
B vmPFC	-	-	-	-	-	10	**36**	0	46	-9

Note. Activations in the conjunctions shown were active above the active baseline within every condition at an individual voxel threshold of $p < .005$ and a cluster threshold of 34 voxels, yielding a whole-brain threshold of $p < .05$ corrected for multiple comparisons. Note that clusters smaller than 34 voxels reflect fragments of larger supra-threshold clusters that fell into the shared overlap across conditions. Abbreviations: **L**: left; **R**: right; **B**: bilateral, **ATL**: anterior temporal lobe, **dlPFC**: dorsolateral prefrontal cortex, **dmPFC**: dorsomedial prefrontal cortex, **MTG**: middle temporal gyrus, **OFC**: orbitofrontal cortex, **PCC**: posterior cingulate cortex, **SMA**: supplementary motor area, **STG**: superior temporal gyrus, **STS**: superior temporal sulcus, **vlPFC**: ventrolateral prefrontal cortex, **vmPFC**: ventromedial prefrontal cortex.

Table 7. Significant activations in conceptual representation pairwise contrasts between all possible pairs of individuals in Phase 2 (I, Relative, Friend, Acquaintance).

Region	I vs. Relative						I vs. Friend						I vs. Acquaintance					
	BA	Volume	Peak t	(x, y, z)			BA	Volume	Peak t	(x, y, z)			BA	Volume	Peak t	(x, y, z)		
L TPJ incl.																		
L supramarginal gyrus,	40,	**603**	6.18	-46	-56	14	40,	**325**	6.47	-52	-38	30	40	**115**	6.58	-56	-38	26
L pSTG,	22,						22,						-	-	-	-	-	-
L pSTS, pMTG	39, 37						39						-	-	-	-	-	-
R TPJ incl.																		
R supramarginal gyrus,	40,	**544**	6.04	58	-38	24	-	-	-	-	-	-	-	-	-	-	-	-
R pSTG,	22, 21																	
R pSTS, pMTG	39, 37																	
B mid cingulate	31, 24	**157**	5.27	2	-14	44	-	-	-	-	-	-	-	-	-	-	-	-
B dACC	32, 24,	**135**	4.23	-4	52	12	-	-	-	-	-	-	32, 24	**80**	4.23	-8	28	26
B dmPFC, frontopolar	10						-	-	-	-	-	-	-	-	-	-	-	-
R extrastriate ctx,	18, 19	**124**	5.56	38	-86	8	-	-	-	-	-	-	-	-	-	-	-	-
R pMTG	39																	
B precuneus	7	**99**	4.62	-2	-46	50	-	-	-	-	-	-	-	-	-	-	-	-
L vlPFC,	44, 45,	**80**	4.82	-46	-2	12	-	-	-	-	-	-	-	-	-	-	-	-
L insula	13																	
R extrastriate ctx.	19	**63**	4.14	28	-70	30	-	-	-	-	-	-	-	-	-	-	-	-
R insula	13	**45**	4.74	40	-8	14	-	-	-	-	-	-	-	-	-	-	-	-
L extrastriate ctx.	19	**39**	4.10	-20	-80	30	-	-	-	-	-	-	-	-	-	-	-	-

Region	Relative vs. I						Relative vs. Friend						Relative vs. Acquaintance					
	BA	Volume	Peak t	(x, y, z)			BA	Volume	Peak t	(x, y, z)			BA	Volume	Peak t	(x, y, z)		
B retrosplenial ctx.	31,23,30	**115**	-5.36	-4	-58	24	-	-	-	-	-	-	-	-	-	-	-	-

Region	Friend vs. I						Friend vs. Relative						Friend vs. Acquaintance					
	BA	Volume	Peak t	(x, y, z)			BA	Volume	Peak t	(x, y, z)			BA	Volume	Peak t	(x, y, z)		
B retrosplenial ctx.	31,23,30	**263**	-6.08	2	-62	24	-	-	-	-	-	-	-	-	-	-	-	-
B sgACC	25	**56**	-6.68	2	10	-4	-	-	-	-	-	-	-	-	-	-	-	-

R supramarginal gyrus,	-	-	-	-	-	-	40,	**266**	-6.13	52	-50	18	-	-	-	-	-	-
R STG	-	-	-	-	-	-	22,42,41											
B mid & dorsal ACC	-	-	-	-	-	-	24,32	**171**	-5.83	-2	-10	36	-	-	-	-	-	-
L dACC	-	-	-	-	-	-	-	-	-	-	-	-	32	**109**	6.03	-14	32	26

	Acquaintance vs. I						**Acquaintance vs. Relative**						**Acquaintance vs. Friend**					
Region	**BA**	**Volume**	**Peak t**	**(x, y, z)**			**BA**	**Volume**	**Peak t**	**(x, y, z)**			**BA**	**Volume**	**Peak t**	**(x, y, z)**		
B PCC,	31, 23,	**352**	-6.46	-2	-44	14		-	-	-	-	-		-	-	-	-	-
B retrosplenial ctx	30, 29																	
L striate	17,	**98**	-5.06	-16	-82	-22		-	-	-	-	-		-	-	-	-	-
& extrastriate ctx.	18																	
B vmPFC	10	**62**	-5.71	8	46	-6		-	-	-	-	-		-	-	-	-	-
L MTG, ATL	21, 38	**39**	-5.33	-50	-2	-12		-	-	-	-	-		-	-	-	-	-
R STG	22, 42	-	-	-	-	-		**196**	-6.33	58	-38	8		-	-	-	-	-
R STG, MTG, insula	22, 21, 13	-	-	-	-	-		**95**	-5.07	44	-10	0		-	-	-	-	-
R angular gyrus	19, 39	-	-	-	-	-		**40**	-5.26	38	-76	38		-	-	-	-	-

Note. Activations in the contrasts shown were active at an individual voxel threshold of $p < .005$ and a cluster threshold of 34 voxels, yielding a whole-brain threshold of $p < .05$ corrected for multiple comparisons. Abbreviations: **L**: left; **R**: right; **B**: bilateral, **ATL**: anterior temporal lobe, **dACC:** dorsal anterior cingulate cortex, **dmPFC**: dorsomedial prefrontal cortex, **MTG**: middle temporal gyrus, **PCC** posterior cingulate cortex, **sgACC:** subgenual anterior cingulate cortex, **STG**: superior temporal gyrus, **STS**: superior temporal sulcus, **TPJ**: temporoparietal junction, **vlPFC**: ventrolateral prefrontal cortex, **vmPFC**: ventromedial prefrontal cortex.

Table 8. Results from the conceptual representation conjunction across the three tasks in Phase 2 (Individuals, Groups, and semantic Prompts), showing activations shared across all three tasks, subsets of two tasks, and unique for each task.

	Individuals * Groups * Prompts		
Region	**BA**	**Volume**	**CM**
L supramarginal gyrus,	40,	**641**	(-39, 13, 22)
L S1, M1, premotor ctx	2, 3, 4, 6		
L dlPFC,	9, 46		
L vlPFC,	44, 45		
L OFC, insula	47, 11, 13		
L extrastriate ctx.	18, 19	**412**	(-22, -88, -5)
B SMA, dACC,	6, 32	**244**	(-4, 10, 49)
L dmPFC	8		
R extrastriate ctx.	18, 19	**251**	(22, -86, -4)
L MTG	22	**32**	(-46, -40, 3)
L fusiform gyrus	37	**9**	(-37, -58, -10)

	Individuals * Groups					Individuals * Prompts					Groups * Prompts				
Region	**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**			**BA**	**Volume**	**CM**		
B SMA	6	**80**	-4	15	50	6L	**68**	-15	-5	50	-	-	-	-	-
L extrastriate ctx.	17, 18	**15**	-22	-99	-5	19	**24**	-39	-82	-10	-	-	-	-	-
L dlPFC, vlPFC,	-	-	-	-	-	9, 44, 45	**208**	-40	24	11	-	-	-	-	-
& OFC, insula						& 47, 13					-	-	-	-	-
L fusiform gyrus	-	-	-	-	-	37, 20	**105**	-41	-43	-10	-	-	-	-	-
R extrastriate ctx.	-	-	-	-	-	18	**69**	29	-89	-8	-	-	-	-	-
R premotor ctx.	-	-	-	-	-	6	**36**	26	-10	50	-	-	-	-	-
L premotor ctx.	-	-	-	-	-	6	**23**	-33	1	45	-	-	-	-	-

		Individuals				**Groups**				**Prompts**			
Region	**BA**	**Volume**	**CM**			**Volume**	**CM**			**Volume**	**CM**		
L supramarginal gyrus,	*many*	-	-	-	-	-	-	-	-	**223**	-53	11	33
L S1/M1,	-	-	-	-		-	-	-	-	*included*			
L dlPFC, vlPFC, OFC,		**626**	-48	10	-3	-	-	-	-	**82**	-35	41	-4
L insula,		*included*				-	-	-	-	-	-	-	-
L premotor ctx.,		-	-	-	-	-	-	-	-	-	-	-	-
L lateral temporal ctx.,		*included*				-	-	-	-	-	-	-	-
L fusiform gyrus		-	-	-	-	-	-	-	-	**64**	-41	-44	-19
L dmPFC		**102**	-8	43	38	-	-	-	-	-	-	-	-
B SMA		**77**	0	5	58	-	-	-	-	-	-	-	-
R extrastriate ctx.	18	**39**	30	-93	5	-	-	-	-	-	-	-	-
L vmPFC, sgACC	10, 11, 32	**234**	-2	41	-7	-	-	-	-	-	-	-	-
L angular gyrus, pMTG	39	**73**	-47	-67	24	-	-	-	-	-	-	-	-
B PCC	31	**69**	-4	-55	24	-	-	-	-	-	-	-	-
R lateral OFC	47	**35**	33	30	-4	-	-	-	-	-	-	-	-
R premotor ctx.	6	-	-	-	-	-	-	-	-	**84**	29	-8	39

Note. Activations in the conjunctions shown were active above the active baseline within every condition at an individual voxel threshold of $p < .005$ and a cluster threshold of 34 voxels, yielding a whole-brain threshold of $p < .05$ corrected for multiple comparisons. Note that clusters smaller than 34 voxels reflect fragments of larger supra-threshold clusters that fell into the shared overlap across conditions. Abbreviations: **L**: left; **R**: right; **B**: bilateral, **dACC**: dorsal anterior cingulate cortex, **dlPFC**: dorsolateral prefrontal cortex, **dmPFC**: dorsomedial prefrontal cortex, **M1**: primary motor cortex, **MTG**: middle temporal gyrus, **OFC**: orbitofrontal cortex, **PCC**: posterior cingulate cortex, **S1**: primary somatosensory cortex, **sgACC**: subgenual anterior cingulate cortex, **SMA**: supplementary motor cortex, **vlPFC**: ventrolateral prefrontal cortex.

Table 9. Significant activations in conceptual representation pairwise contrasts between all possible pairs of tasks in Phase 2 (Individuals, Groups, and semantic Prompts).

BA	Region	Volume	Peak t	Peak voxel		
A) Individual properties > Group properties						
45,	**L** vlPFC (pars triangularis),	**366**	5.26	-56	22	20
47,	**L** lateral OFC (pars orbitalis),					
13,	**L** anterior insula,					
38	**L** anterior temporal lobe					
10, 11,	**B** vmPFC,	**313**	6.57	-2	32	0
24, 32	**B** sgACC					
47	**R** lateral OFC (pars orbitalis)	**53**	5.83	34	32	-6
B) Individual properties > Prompted properties						
9, 10, 11	**L** dmPFC, **B** vmPFC,	**375**	6.34	-2	46	-10
24, 32	**B** sgACC					
45,	**L** vlPFC (pars triangularis),	**283**	6.17	-44	14	-22
47,	**L** lateral OFC (pars orbitalis),					
38,	**L** anterior temporal lobe					
31, 23	**L** PCC	**182**	4.52	-14	-52	8
47	**L** lateral OFC (pars orbitalis)	**39**	4.30	-28	16	-10
C) Group properties > Prompted properties						
	none					
D) Group properties > Individual properties						
	none					
E) Prompted properties > Individual properties						
9, 44	**R** dlPFC, vlPFC (pars opercularis)	**37**	-4.41	50	4	26
F) Prompted properties > Group properties						
9, 46, 10	**L** dlPFC	**122**	-5.40	-34	20	30
47	**L** OFC					
37, 20	**L** fusiform gyrus, MTG	**106**	-4.96	-46	-38	-10
NA	**B** thalamus (MGN, I think)	**51**	-4.43	-4	-26	-4
6	**R** premotor ctx.	**51**	-5.17	28	-8	50
6	**L** premotor ctx.	**43**	-4.61	-20	4	50
10	**R** frontopolar ctx., mOFC	**35**	-4.58	22	38	0
9	**R** dlPFC	**34**	-4.38	56	10	32

Note. Activations in the contrasts shown were active at an individual voxel threshold of $p < .005$ and a cluster threshold of 34 voxels, yielding a whole-brain threshold of $p < .05$ corrected for multiple comparisons. Abbreviations: **L**: left; **R**: right; **B**: bilateral, **dlPFC**: dorsolateral prefrontal cortex, **dmPFC**: dorsomedial prefrontal cortex, **OFC**: orbitofrontal cortex, **PCC**: posterior cingulate cortex, **sgACC**: subgenual anterior cingulate cortex, **vlPFC**: ventrolateral prefrontal cortex, **vmPFC**: ventromedial prefrontal cortex.

Table 10. Significant activations in pairwise contrasts between core self (I) and core representations for others (Relative, Friend, and Acquaintance).

BA	Region	Volume	Peak t	Peak voxel		
Self > Other						
37, 19	**R** extrastriate ctx., fusiform g.	474	9.00	38	-52	-6.5
37, 19	**L** extrastriate ctx., fusiform g.	258	6.28	-44	-64	0
NA	**L** caudate, putamen, & globus pallidus (ext.)	84	4.77	-20	-8	8
40	**R** TPJ (supramarginal gyrus)	67	4.09	44	-32	26
40, 42, 22	**L** TPJ (supramarginal gyrus) & STG	58	4.51	-44	-38	6
Other > Self						
17, 18, 19, 31	**B** striate & extrastriate ctx., **L** PCC	1770	-10.95	-16	-86	-10

Note. Activations in the contrasts shown were active at an individual voxel threshold of $p < .005$ and a cluster threshold of 34 voxels, yielding a whole-brain threshold of $p < .05$ corrected for multiple comparisons. Abbreviations: **L**: left; **R**: right; **B**: bilateral, **PCC**: posterior cingulate cortex, **TPJ**: temporoparietal junction.

Table 11. Significant activations in pairwise contrasts between conceptual self (I) and conceptual representations for others (Relative, Friend, and Acquaintance).

BA	Region	Volume	Peak t	Peak voxel		
Self > Other						
40, 39, 22, 21	**L** TPJ (supramarginal gyrus), **L** pSTG, pSTS, pMTG	429	6.88	-52	-34	30
32, 24, 10	**B** dACC, dmPFC	85	5.02	-8	38	12
19	**R** extrastriate ctx.	49	4.40	38	-88	8
39, 19	**R** posterior MTG	45	5.05	46	-76	20
8	**L** dlPFC	39	5.03	-16	44	50
Other > Self						
31, 23, 30, 29	**R** PCC, retrosplenial ctx.,	317	-6.25	2	-62	24
18	**L** extrastriate ctx.	81	-4.78	-16	-82	-16
10	**B** vmPFC	75	-5.88	8	46	-6
18	**R** extrastriate ctx.	38	-4.70	16	-80	-16
NA	**B** caudate nucleus	37	-6.63	-2	2	0

Note. Activations in the contrasts shown were active at an individual voxel threshold of $p < .005$ and a cluster threshold of 34 voxels, yielding a whole-brain threshold of $p < .05$ corrected for multiple comparisons. Abbreviations: **L**: left; **R**: right; **B**: bilateral, **dACC**: dorsal anterior cingulate cortex, **dmPFC**: dorsomedial prefrontal cortex, **MTG**: middle temporal gyrus, **PCC**: posterior cingulate cortex, **STG**: superior temporal gyrus, **STS**: superior temporal sulcus, **TPJ**: temporoparietal junction, **vmPFC**: ventromedial prefrontal cortex.

Figures

Figure 1. Core and conceptual selves. We dichotomize representations of the self (and of other individuals) according to distinctions drawn in neuroscience and philosophy (Epstein, 1973; Neisser, 1988; Damasio, 1999; Gallagher, 2000; Legrand, 2007)

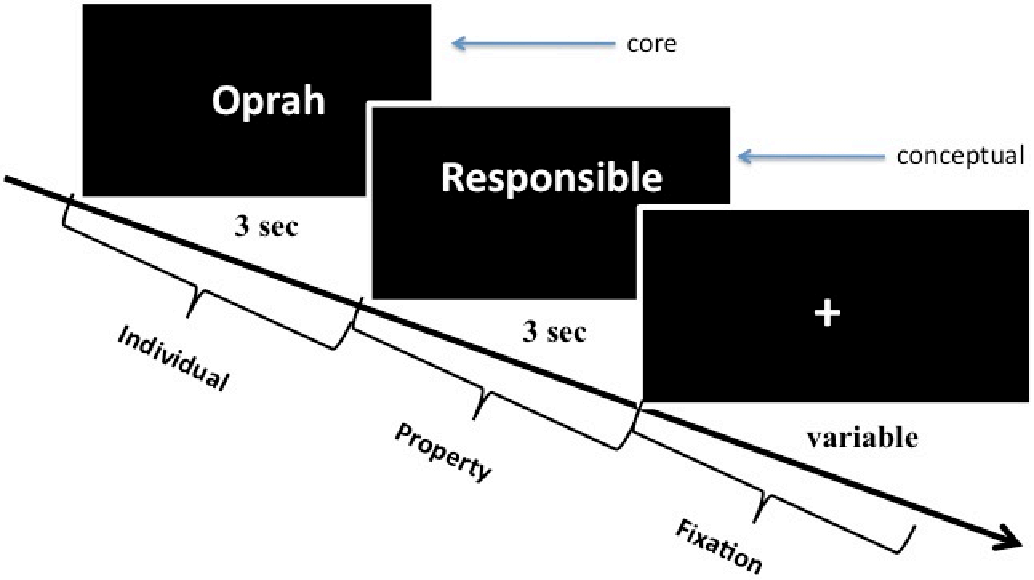


Figure 2. The structure of a trial. An *individual* ("Oprah", in this case) is presented for 3 seconds. During this phase, we expected that a core representation was generated for that individual. A *property* ("Responsible", in this case) is presented for 3 seconds (except during catch trials), and participants make a forced response. During this phase, we expected that an extended representation was formed by binding conceptual information to the minimal representation just created. Finally, during fixation, participants perform the auditory Baseline Task wherein they press a button upon hearing a tone.

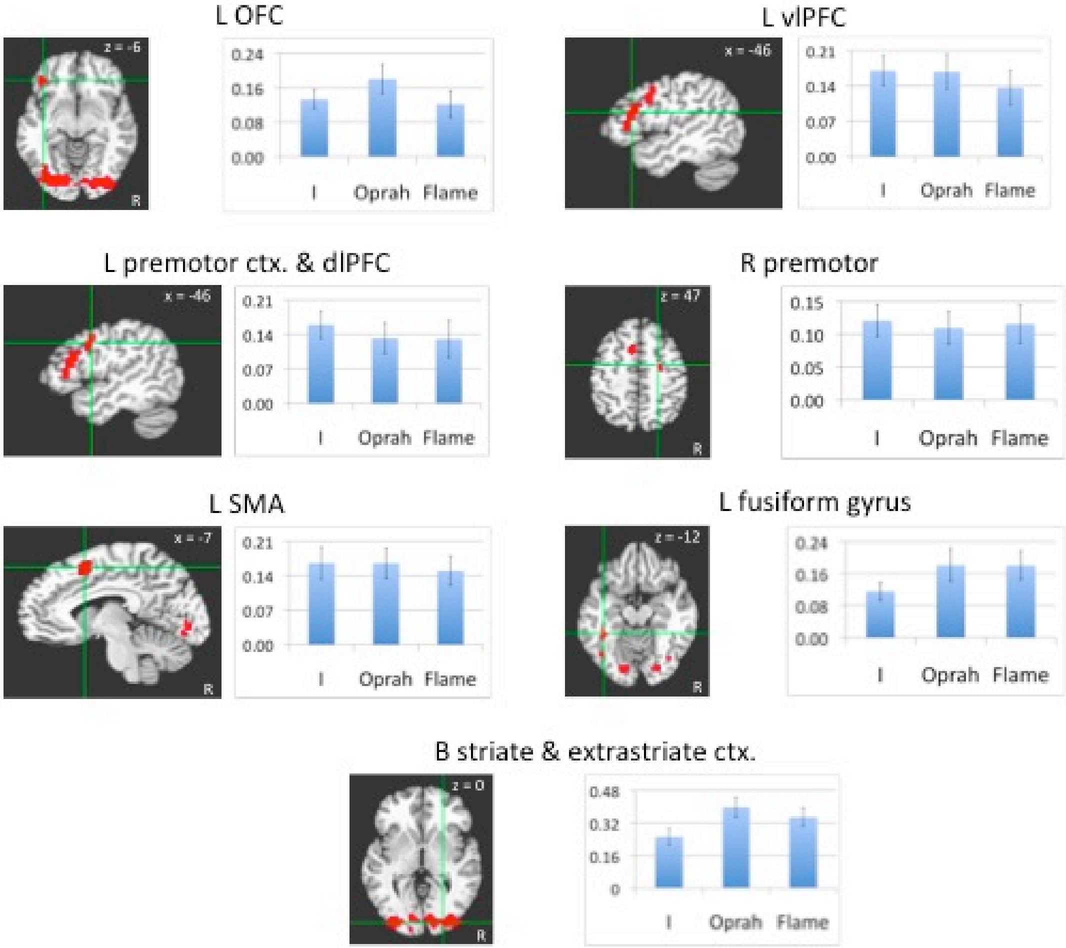
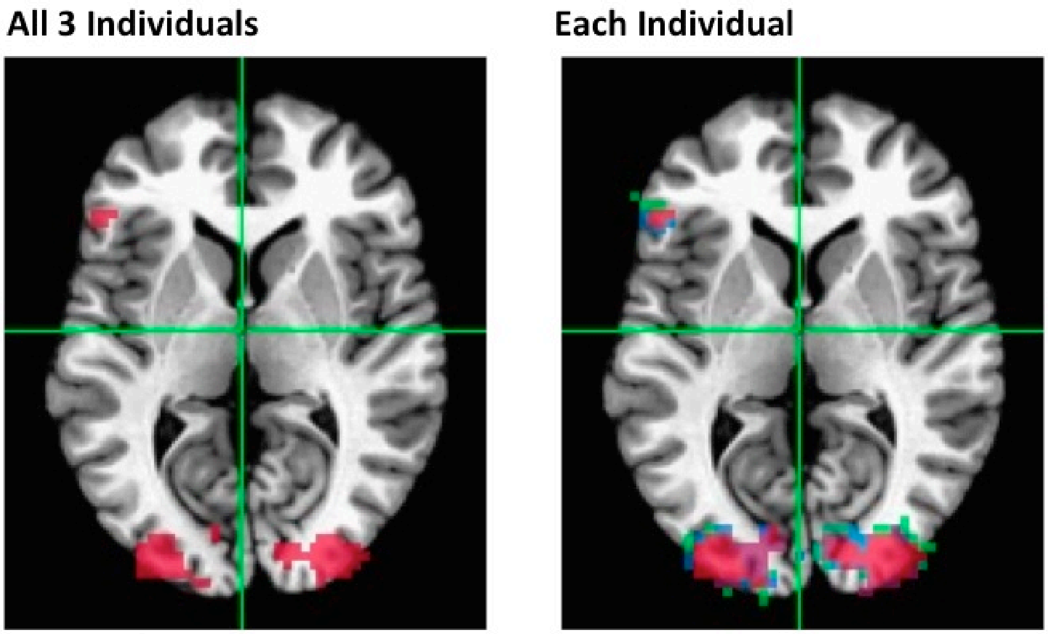


Figure 3. The Core Self Network. Brain regions active for the conjunction of all three *individual* conditions, that is, {I, Oprah, Flame}. We call this set of regions the "identity network (IDN)", and propose that it instantiates minimal representations of individuals such as the self, other people, and objects (tokens). Clusters were thresholded at 27 voxels to achieve a brainwise α = .05



All 3 Individuals

Each Individual

Figure 4. Substantial overlap between individuals. Left: activations from the intersection of all three individuals. Right: union of all three individuals. Note that the activations differ only in extent, not in location.

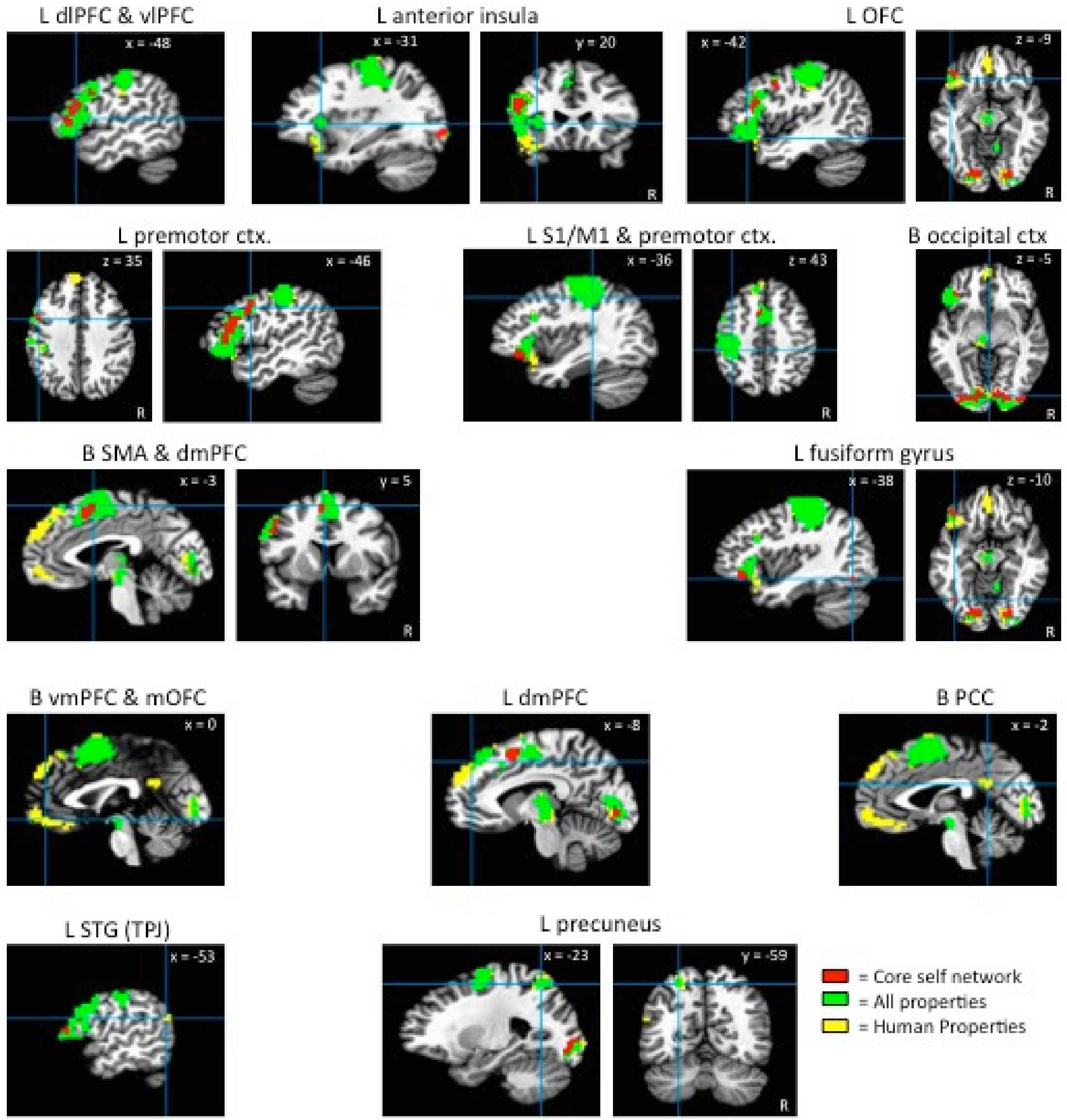
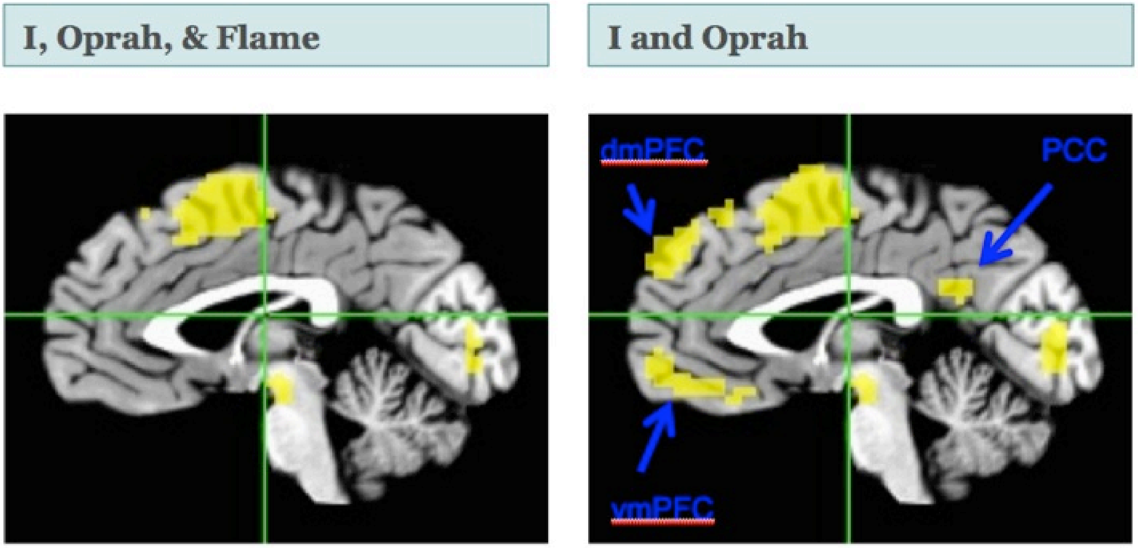


Figure 5. Activations across property conditions. Red: the CSN from Figure 1. Green: brain regions active for the conjunction of all *property* conditions, that is, {Abstract, Concrete} X {I, Oprah, Flame}. Yellow: brain regions active for the conjunction of the two *property* conditions for humans, that is, {Abstract, Concrete} X {I, Oprah}. Note that 1) the IDN is active in both the *individual* and *property* conditions, and that 2) the DMN is active only in the *property* conditions for humans. Clusters were thresholded at 27 voxels to achieve a brainwise α = .05


I, Oprah, & Flame

I and Oprah

Figure 6. Activations across property conditions. Red: the IDN from Figure 1. Green: brain regions active for the conjunction of all *property* conditions, that is, {Abstract, Concrete} X {I, Oprah, Flame}. Yellow: brain regions active for the conjunction of the two *property* conditions for humans, that is, {Abstract, Concrete} X {I, Oprah}. Note that 1) the IDN is active in both the *individual* and *property* conditions, and that 2) the DMN is active only in the *property* conditions for humans. Clusters were thresholded at 27 voxels to achieve a brainwise $\alpha = .05$

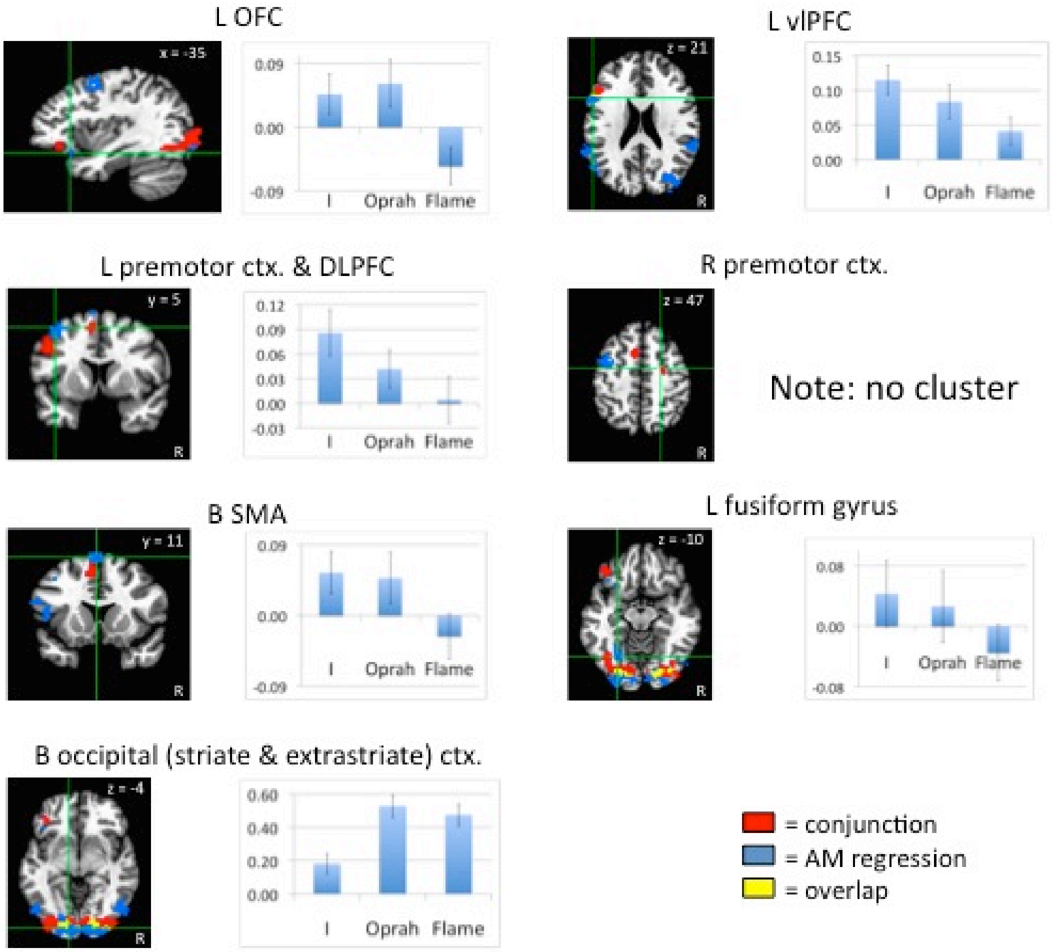


Figure 7. Correlations with self-relatedness in the *individual* conditions. Red: the CSN from Figure 01. Blue: activations positively correlated with self-relatedness in the individual conditions (except for bilateral occipital cortex, which was negatively correlated) . The self-relatedness regressor was coded as follows: {I = 3; Oprah = 2; Flame = 1}. Yellow: overlap between the two. Note that many of the clusters are adjacent but overlap is minimal. Clusters were thresholded at 27 voxels to achieve a brainwise $\alpha = .05$.

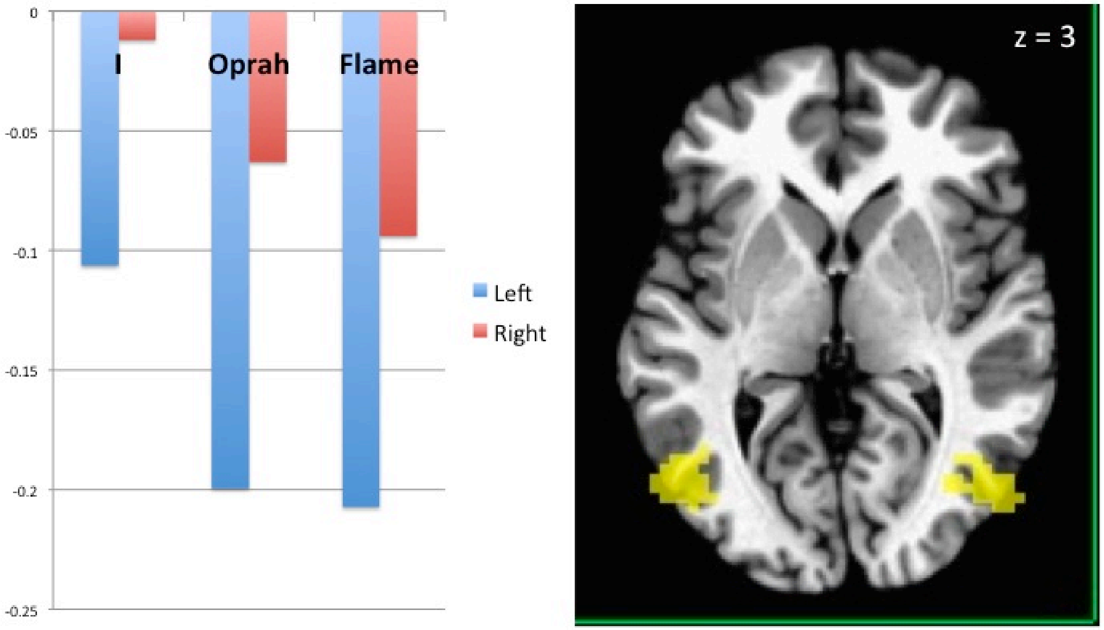
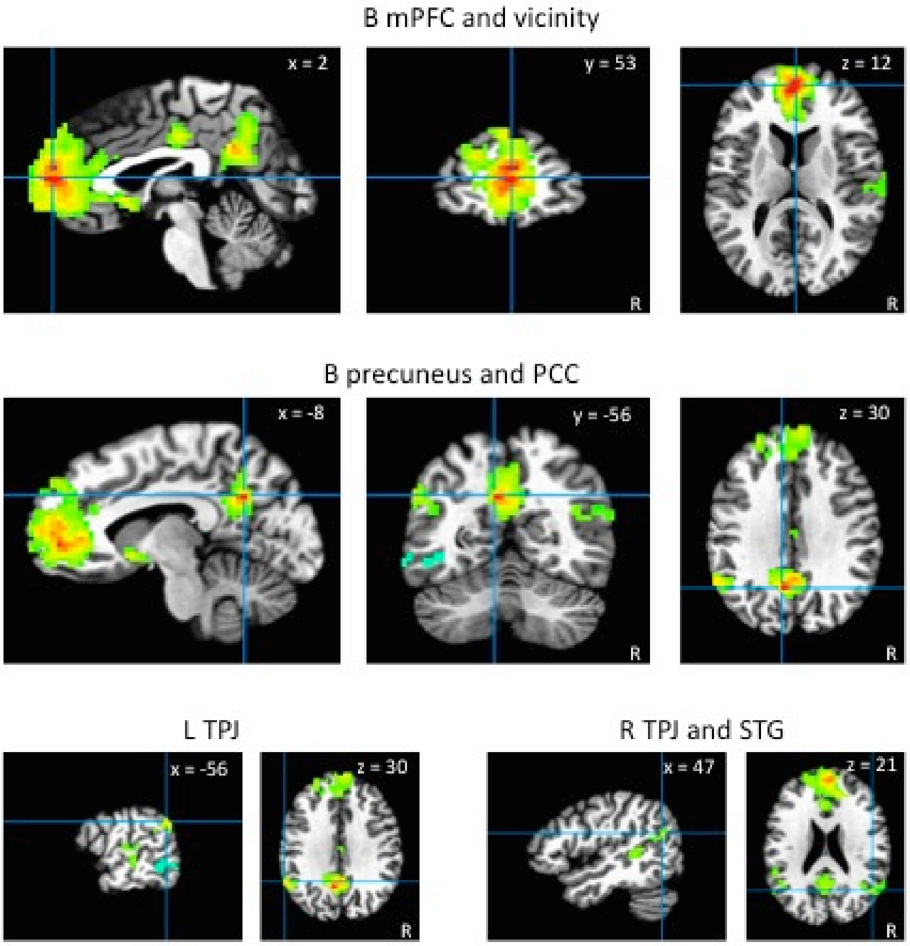


Figure 8. Activation in the extrastriate body area (EBA) correlated positively with self-relatedness.



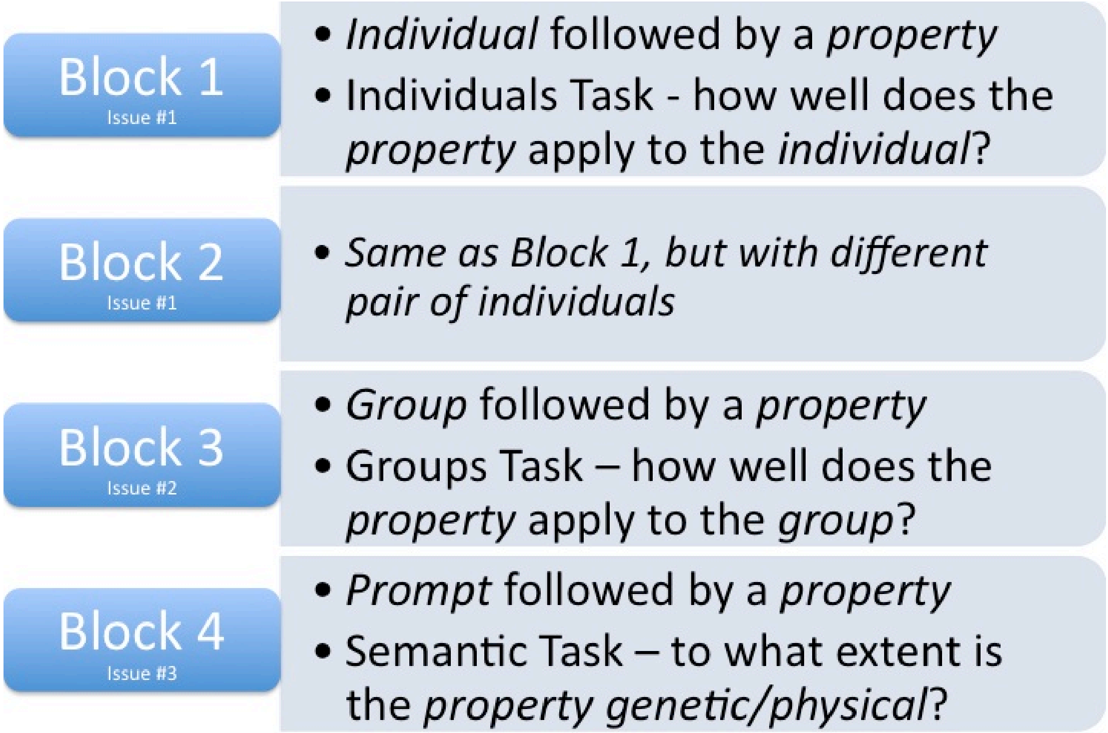
B mPFC and vicinity

B precuneus and PCC

L TPJ

R TPJ and STG

Figure 9. Activations positively correlated with self-relatedness in the *property* conditions. The self-relatedness regressor was coded as follows: {I = 0; Oprah = 1; Flame = 3}. Clusters were thresholded at 27 voxels to achieve a brainwise $\alpha = .05$. Note the similarities between the default mode network and the regions identified here.



Block 1
Issue #1

- *Individual* followed by a *property*
- Individuals Task - how well does the *property* apply to the *individual*?

Block 2
Issue #1

- *Same as Block 1, but with different pair of individuals*

Block 3
Issue #2

- *Group* followed by a *property*
- Groups Task – how well does the *property* apply to the *group*?

Block 4
Issue #3

- *Prompt* followed by a *property*
- Semantic Task – to what extent is the *property genetic/physical*?

Figure 10. Four blocks present in each run. In blocks 1 and 2, participants rate how well *properties* apply to *individuals*. In block 3, participants rate how well *properties* apply to *groups*. In block 4, participants are given *properties* after one of two semantic prompts ("Genetic" or "Physical"), and are asked to rate to what extent the *property* is genetic/physical. In addition to replicating previous results, each block is designed to address one of three Issues (see text).

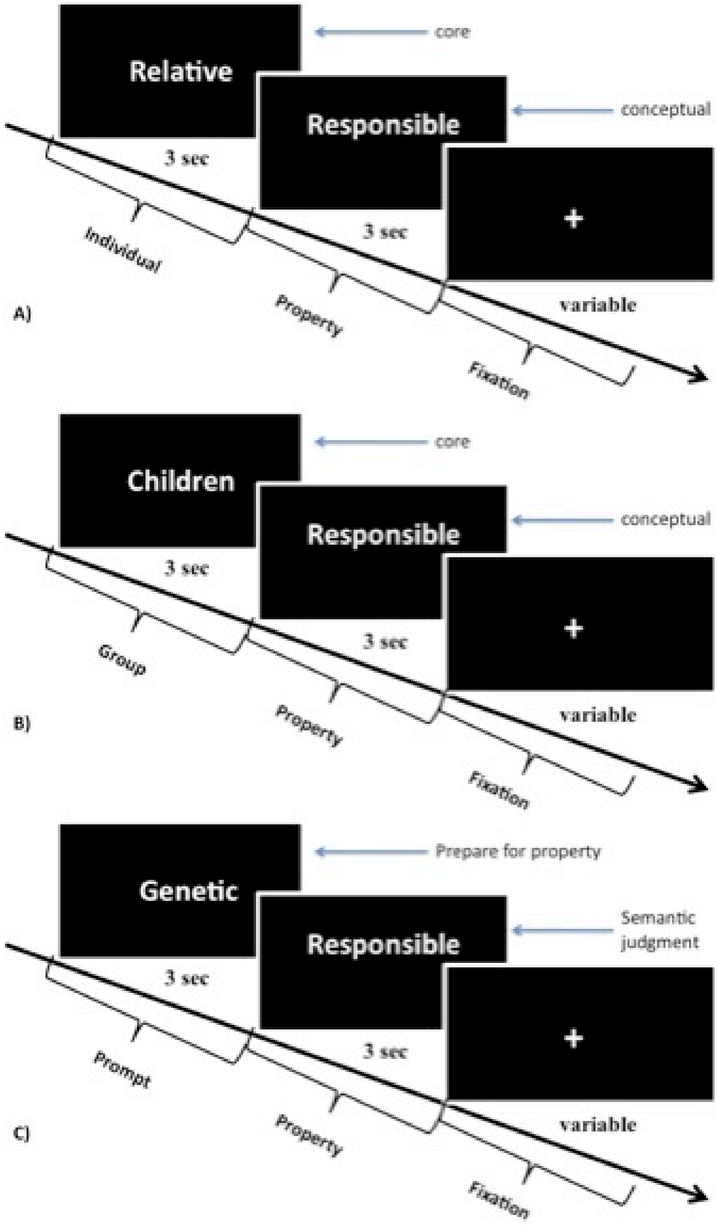


Figure 11. The structure of trials in the current experiment. A) The individuals task. An *individual* ("Relative", in this case) was presented for 3 seconds. During this phase, a core representation was generated for that individual. Next, a *property* ("Responsible", in this case) was presented for 3 seconds (except during catch trials), and participants made a forced response. During this phase, a conceptual representation was formed by binding conceptual information to the core representation just created. Finally, during fixation, participants performed the auditory baseline task wherein they pressed a button upon hearing a tone. B) The groups task was identical to the individuals task, except that a *group* ("Children" in this case) was used instead of an *individual*. C) The semantic task. A semantic *prompt* ("Genetic", in this case) was presented. Next, a *property* ("Responsible", in this case) was presented for 3 seconds (except during catch trials), and participants made a forced response to the question indicated by the prompt. During this phase, we properties were represented per se, as opposed to being predicated of individuals.

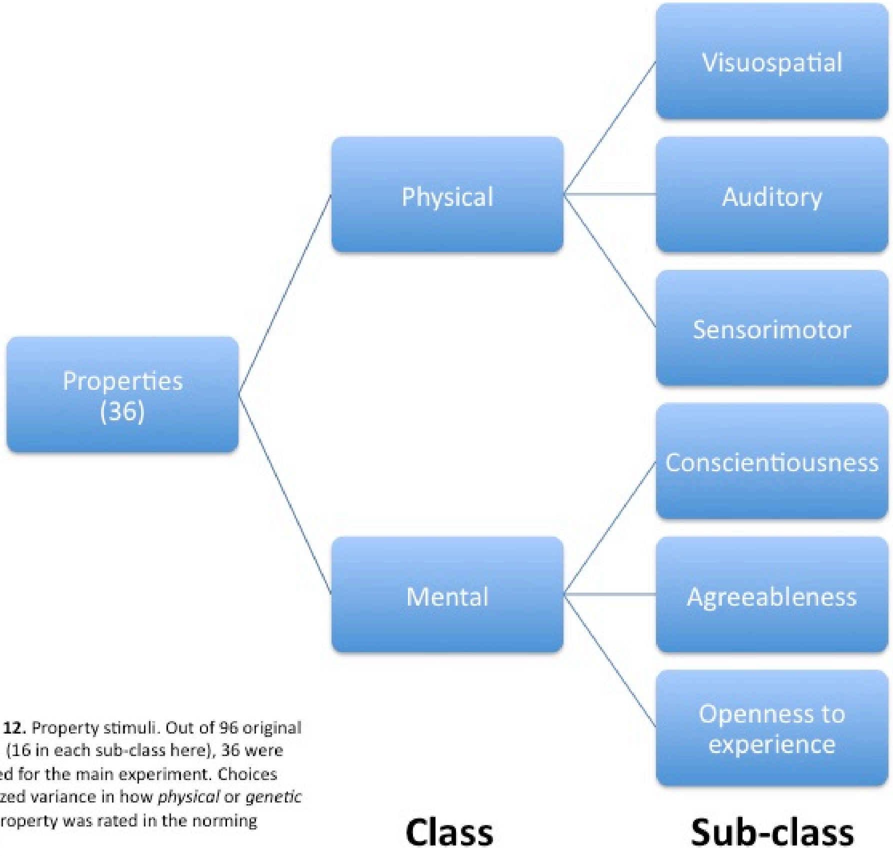


Figure 12. Property stimuli. Out of 96 original stimuli (16 in each sub-class here), 36 were selected for the main experiment. Choices optimized variance in how *physical* or *genetic* each property was rated in the norming study.

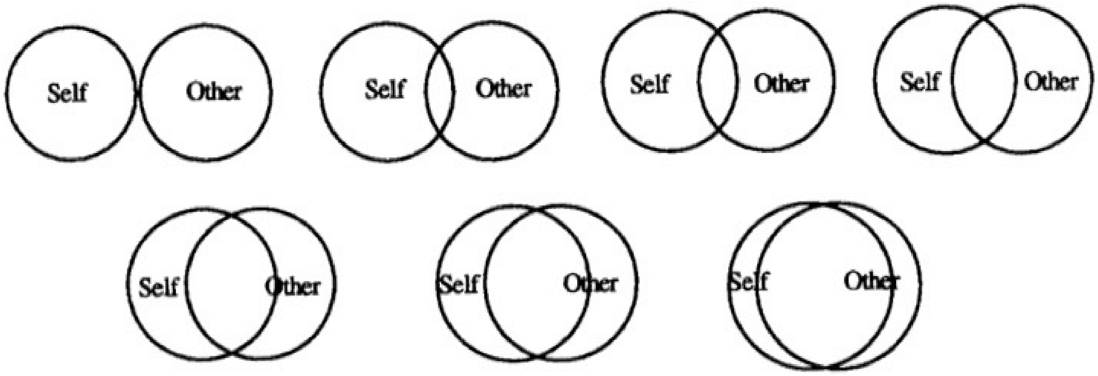

Please circle the picture below which best describes your relationship

Figure 13. The Inclusion of Other in the Self Scale (IOS). Reprinted from Aron, Aron, & Smollan (1992).

Table #	Level of analysis	Core or Conceptual?	Analysis type
1	Person	Core	Conjunction
2	Person	Core	Contrasts
3	Person	Both	AM Regression (Closeness: IOS)
4	Task	Core	Conjunction
5	Task	Core	Contrasts
6	Person	Conceptual	Conjunction
7	Person	Conceptual	Contrasts
8	Task	Conceptual	Conjunction
9	Task	Conceptual	Contrasts
10	Person (Self vs. Others)	Core	Contrasts
11	Person (Self vs. Others)	Conceptual	Contrasts

Figure 14. Organization of the data tables. "Person" level analyses involve I, Relative, Friend, & Acquaintance. "Task" level analyses involve Individuals, Groups, and Prompts

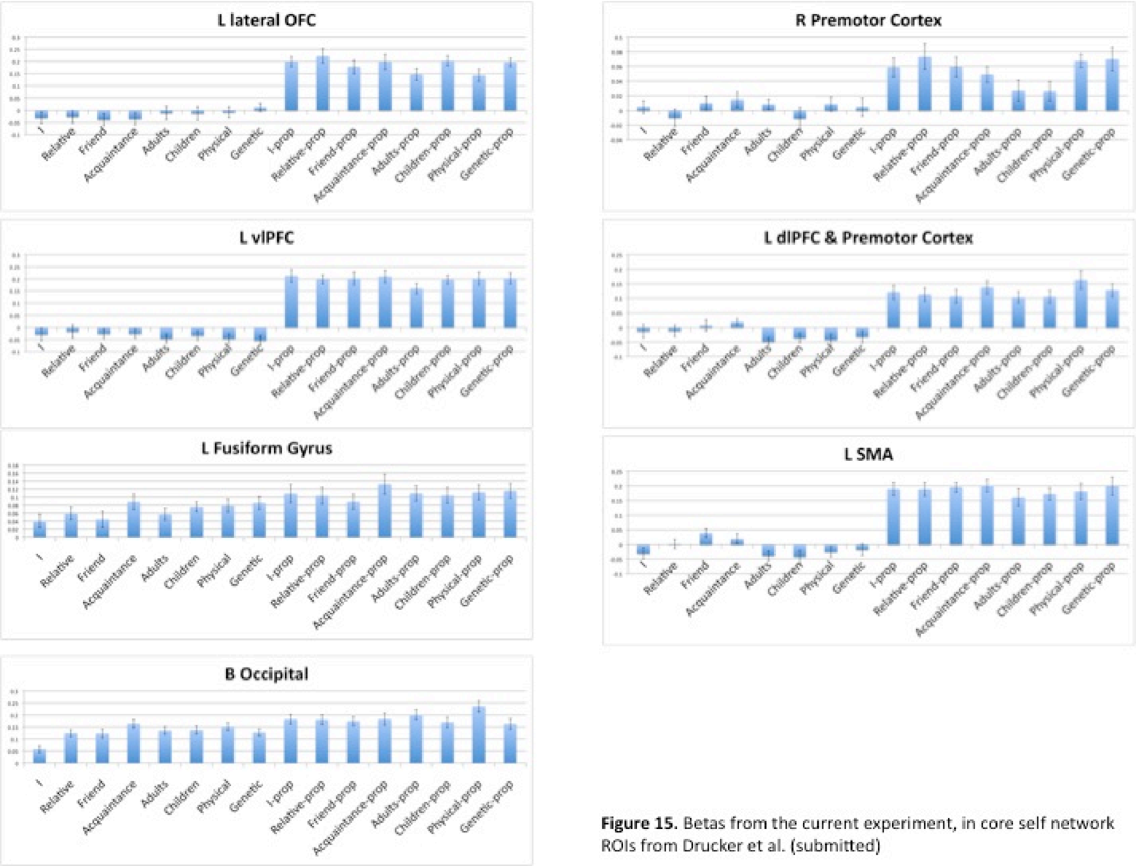


Figure 15. Betas from the current experiment, in core self network ROIs from Drucker et al. (submitted)

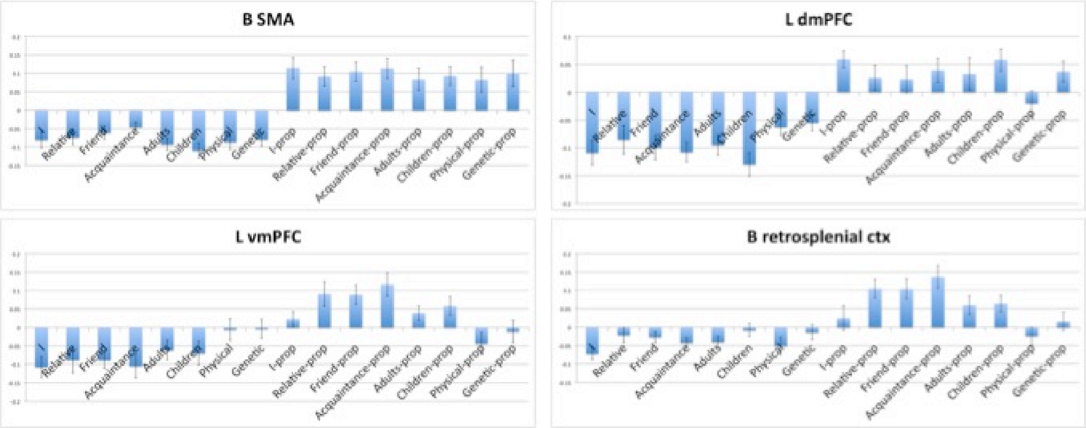
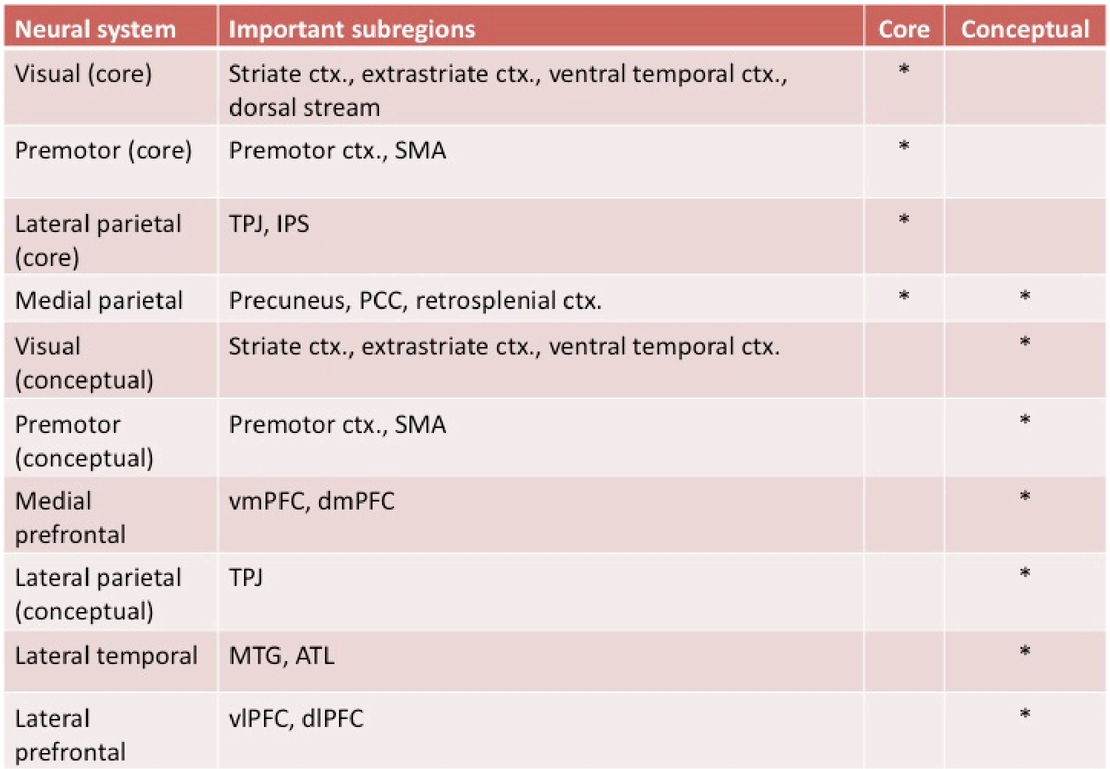
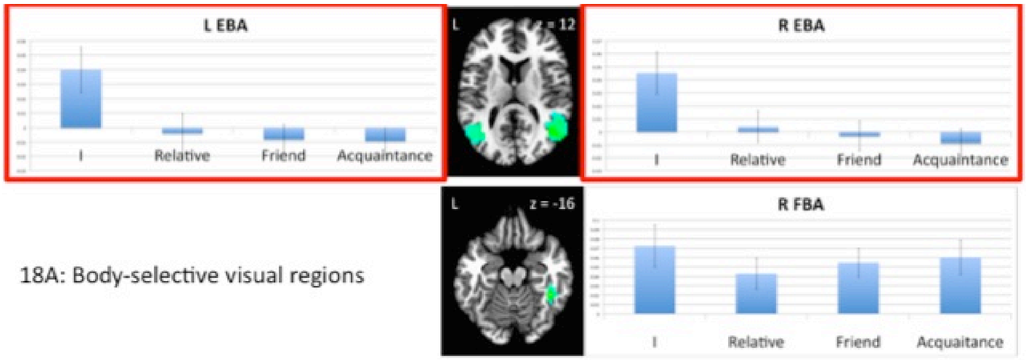


Figure 16. Betas from the current experiment, in conceptual self network ROIs from Drucker et al. (submitted)

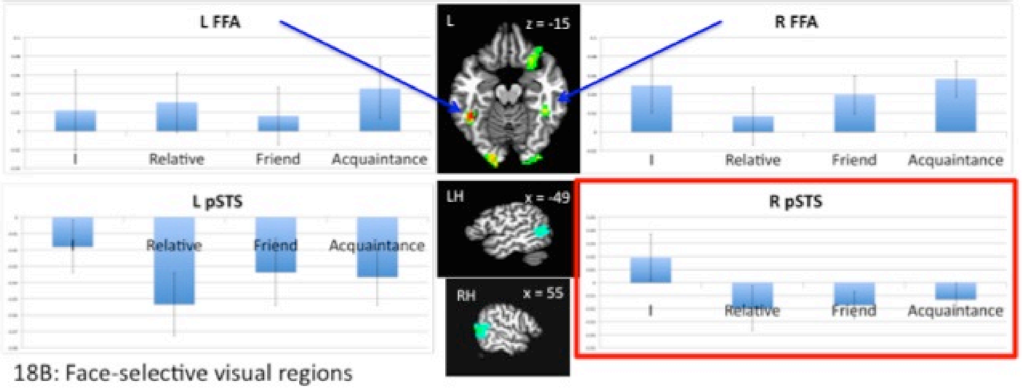


Neural system	Important subregions	Core	Conceptual
Visual (core)	Striate ctx., extrastriate ctx., ventral temporal ctx., dorsal stream	*	
Premotor (core)	Premotor ctx., SMA	*	
Lateral parietal (core)	TPJ, IPS	*	
Medial parietal	Precuneus, PCC, retrosplenial ctx.	*	*
Visual (conceptual)	Striate ctx., extrastriate ctx., ventral temporal ctx.		*
Premotor (conceptual)	Premotor ctx., SMA		*
Medial prefrontal	vmPFC, dmPFC		*
Lateral parietal (conceptual)	TPJ		*
Lateral temporal	MTG, ATL		*
Lateral prefrontal	vlPFC, dlPFC		*

Figure 17. Layout of the first part of the discussion. We discuss each of these neural systems in turn, describing how each contributes to core representations, conceptual representations, or both. The remainder of the discussion deals with the issues that were raised in the introduction.



18A: Body-selective visual regions



18B: Face-selective visual regions

Figure 18. Core representation in ROIs independently localized for visual processing of bodies and faces. Bar graphs represent average GLM coefficients for Phase 1 conditions within each ROI. Error bars represent SEM across participants. Outlined bar graphs indicate core self significantly above every one of the other three conditions.

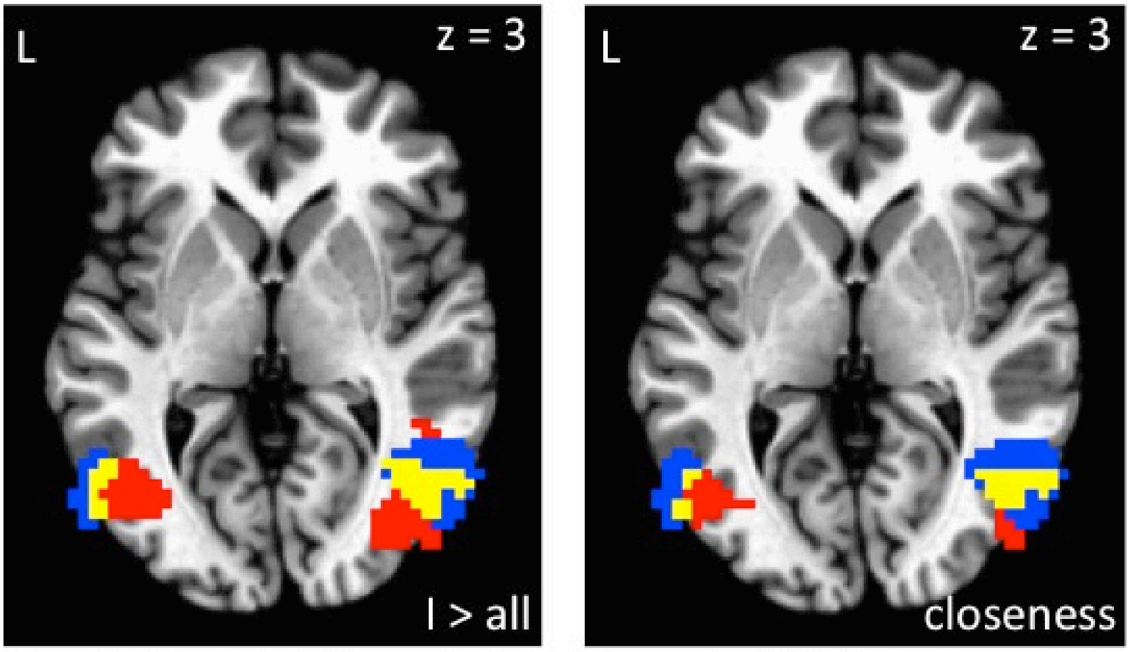


Figure 19. Left: Overlap (yellow) between localizer for visual processing of bodies (blue) and a contrast (red) between core representations for self vs. the other three persons (Relative, Friend, and Acquaintance). Right: Same, but red here is the positive linear association between BOLD and closeness as measured by the IOS.

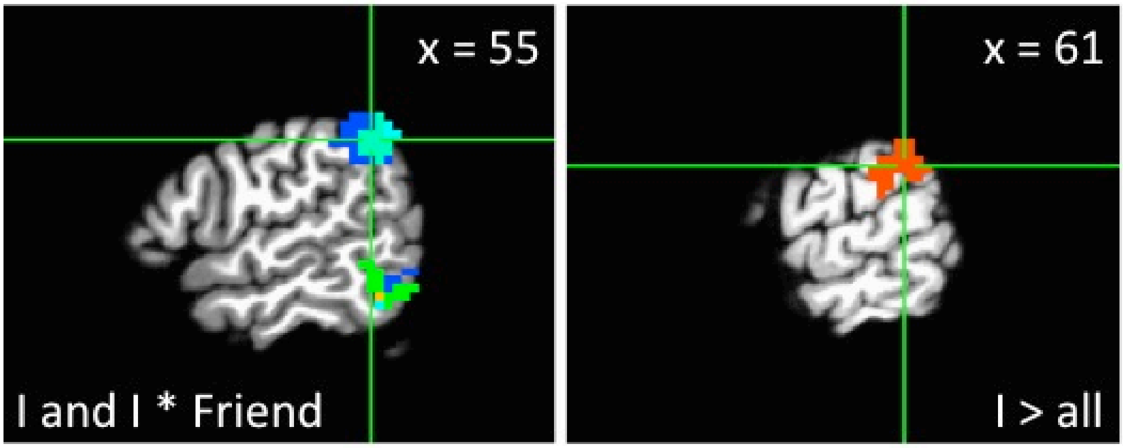


Figure 20. Right temporoparietal junction (TPJ). Left: activation above baseline for I_1 (blue), and I_1 together with Friend_1 (cyan). Right: activation for core self vs. core representations for others.

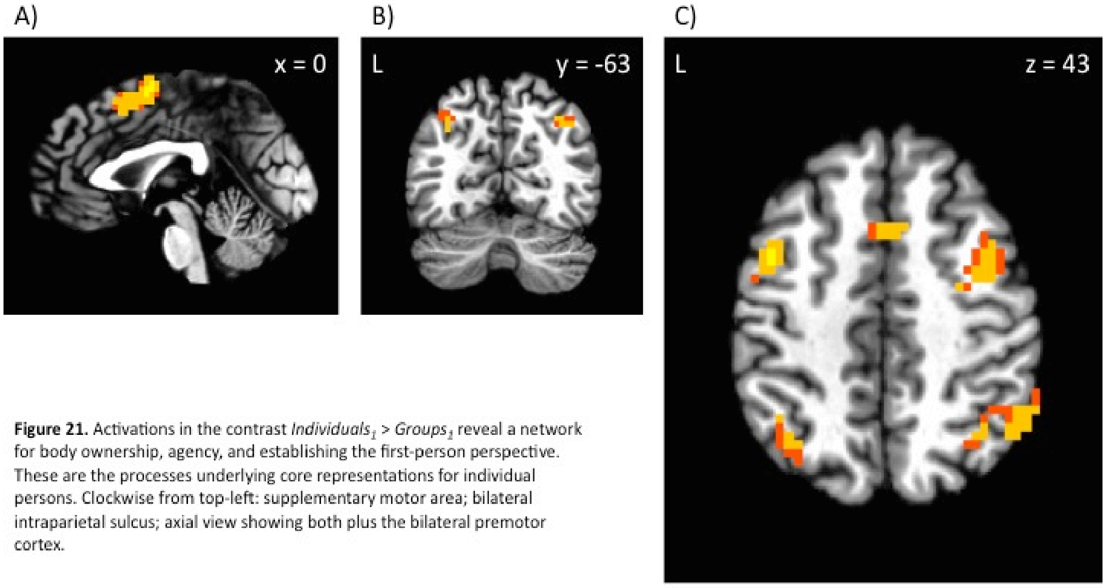


Figure 21. Activations in the contrast *Individuals*$_1$ > *Groups*$_1$ reveal a network for body ownership, agency, and establishing the first-person perspective. These are the processes underlying core representations for individual persons. Clockwise from top-left: supplementary motor area; bilateral intraparietal sulcus; axial view showing both plus the bilateral premotor cortex.

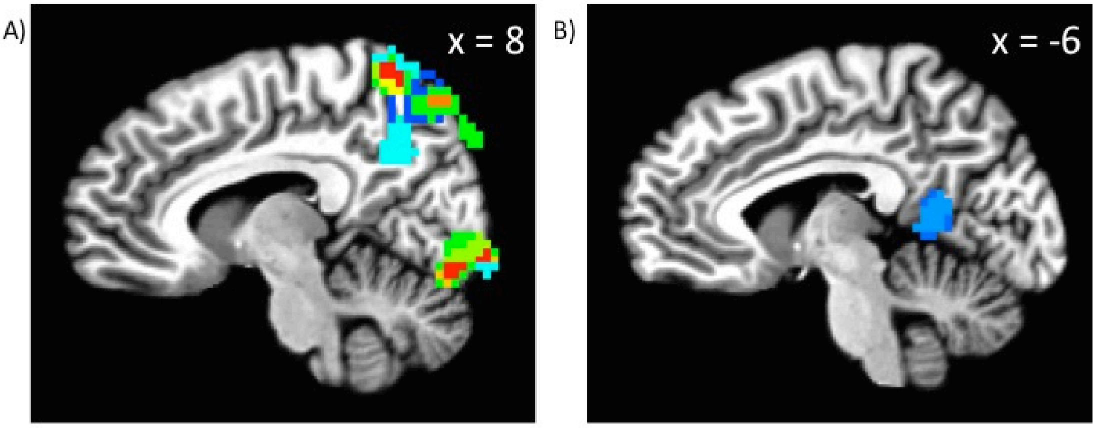


Figure 22. Medial parietal activations during Phase 1. A) Core representations for individual persons had unique activation profiles in right precuneus and PCC. Different colors represent different combinations of persons from the conjunction. Overlap was minimal - red voxels are among only nine voxels significant for all four conditions. B) The left retrosplenial cortex was preferential for groups versus individuals, one of the few clusters significant in that direction

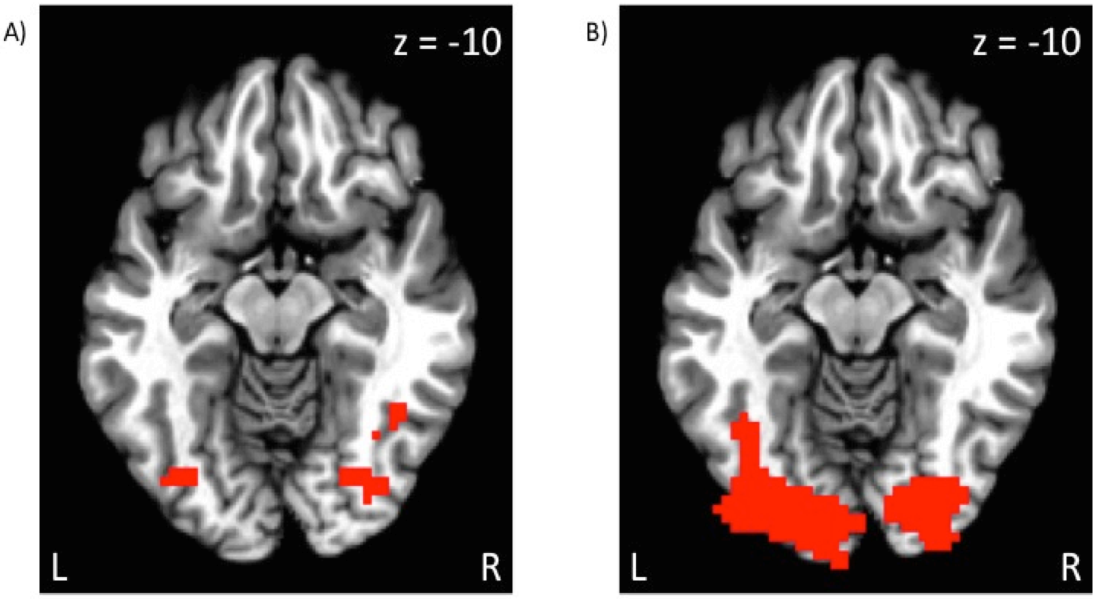


Figure 23. Conjunctions across all four individuals in A) Phase 1 and B) Phase 2. Visual activity was more robust in Phase 2, likely due to conceptual representations being grounded in rich perceptual simulations.

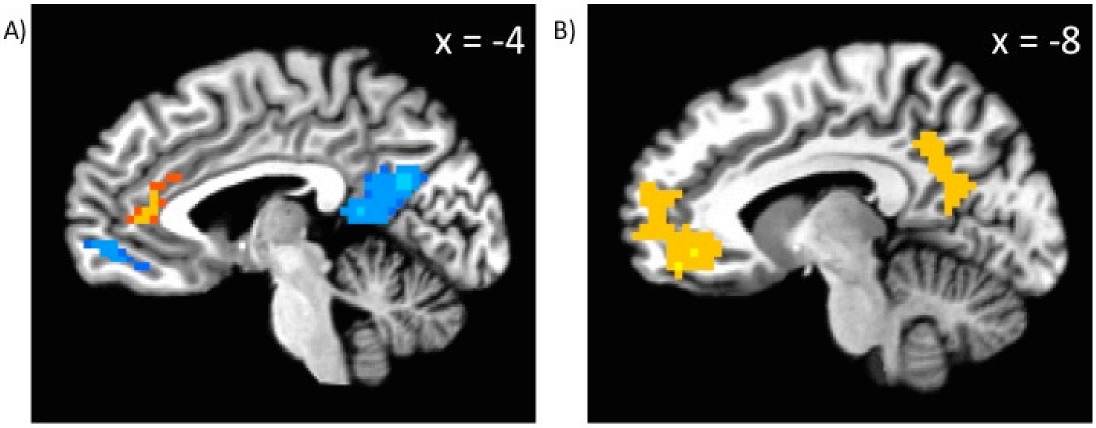


Figure 24. Cortical midline activations during Phase 2. A) Conceptual representations self (orange) versus other persons (blue) activated dorsal and ventral aspects of the left mPFC, respectively. Retrosplenial cortex, preferential for others, is also visible. B) Cortical midline structures were strongly preferential for conceptual representations for individual persons. The contrast *Individuals*$_2$ > *Prompts*$_2$ revels the dmPFC, vmPFC, and medial parietal cortex.

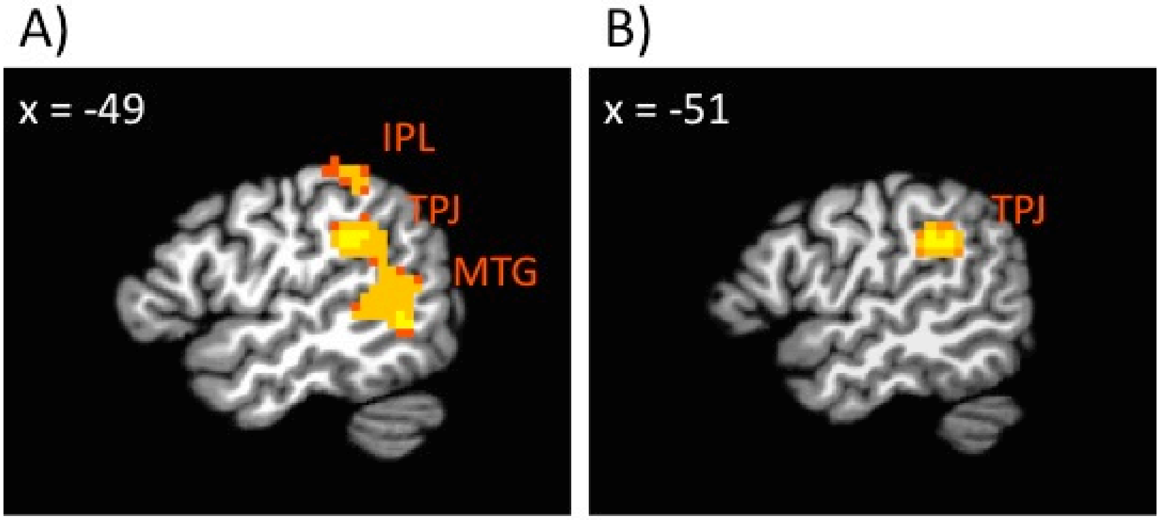


Figure 25. A) Left temporoparietal junction (TPJ) activation for conceptual self vs. conceptual representations for other persons (Relative, Friend, & Acquaintance). Activation extends beyond the TPJ dorsally into the inferior parietal lobule (IPL) and ventrally into the superior and middle temporal gyri (STG & MTG). See Table 2 for differences between specific individuals. **B)** Only the left TPJ was also positively correlated with closeness in Phase 2

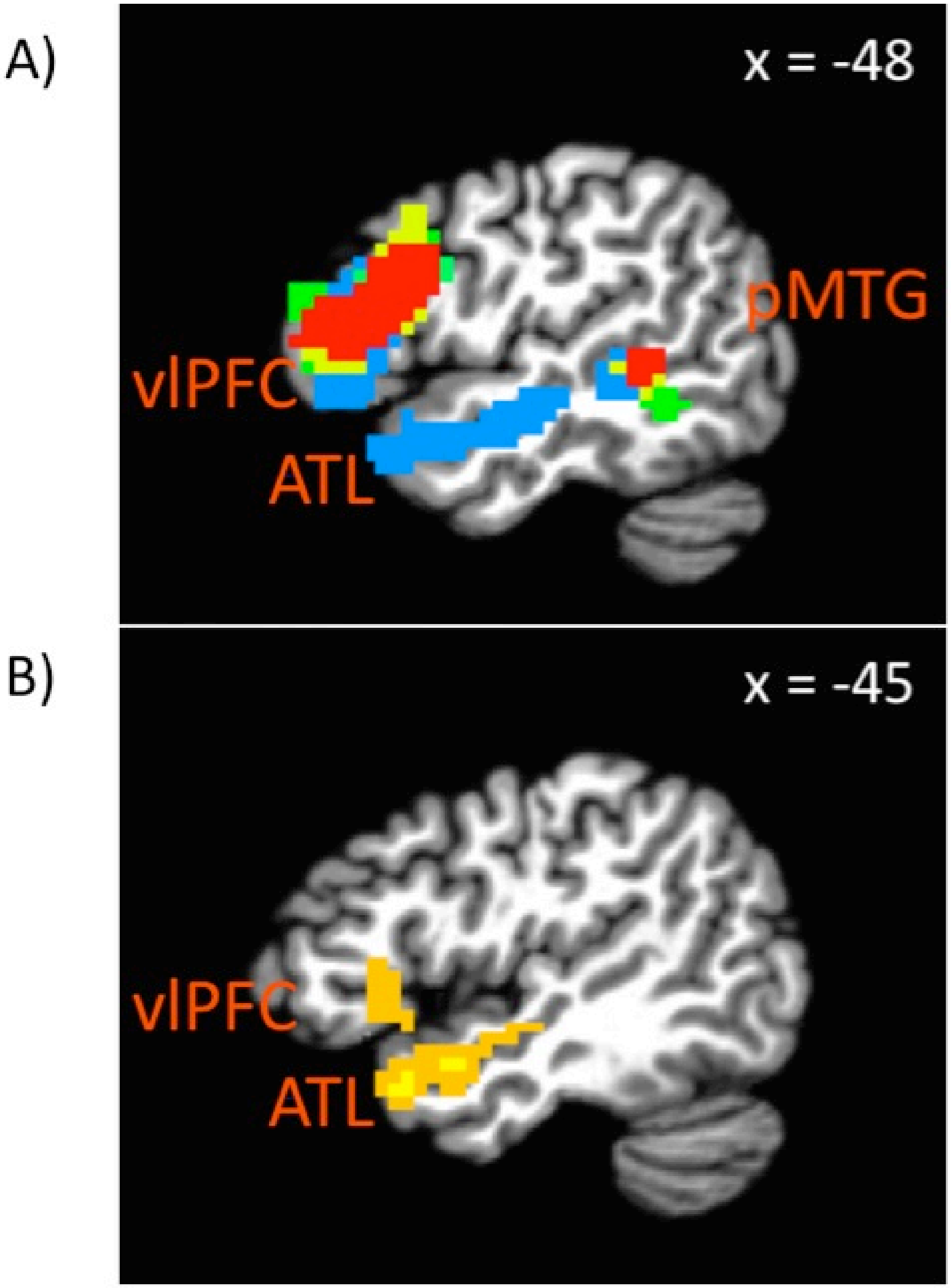


Figure 26. A) Activation against baseline during Phase 2 for *Individuals*$_2$ (blue), *Prompts*$_2$ (green), the 2-way conjunction of *Individuals*$_2$ and *Prompts*$_2$ (yellow), and the 3-way conjunction of *Individuals*$_2$, *Prompts*$_2$, and *Groups*$_2$ (red). Note that activation for *Individuals*$_2$ is far more extensive, continuing anteriorly into the ATL. Activation in the posterior MTG and in the vlPFC is largely the same across all three tasks. B) Activation for *Individuals*$_2$ > *Prompts*$_2$ reveals areas specific fro the conceptual processing of persons in vlPFC and ATL.

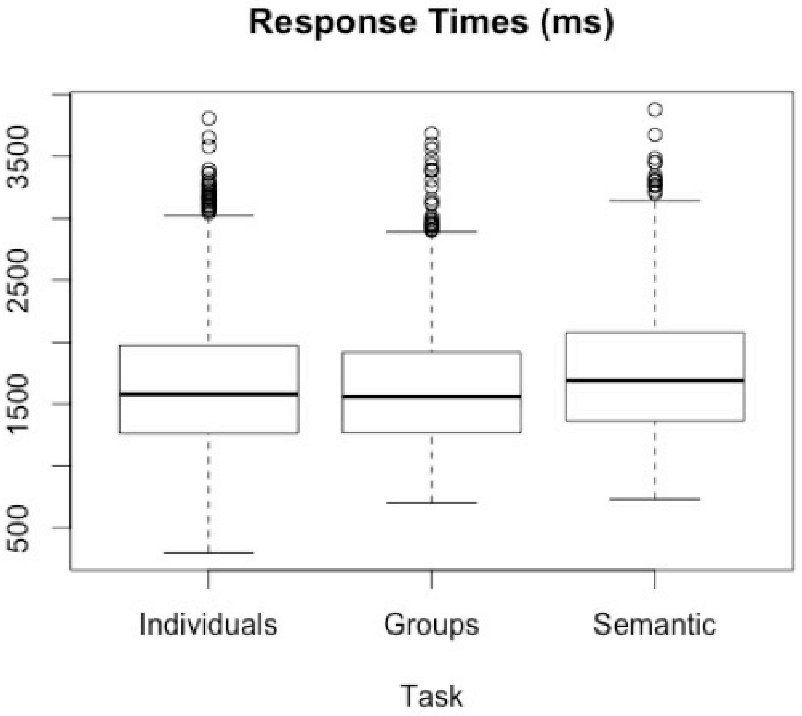


Figure 27. Response latency times for the three tasks. Individuals task: median = 1580 ms, SD = 525. Groups task: median = 1558 ms, SD = 500. Semantic task: median = 1691 ms, SD = 525. Individuals and Groups both differed from Semantic ($p < 10^{-6}$) but not from each other ($p > .05$).

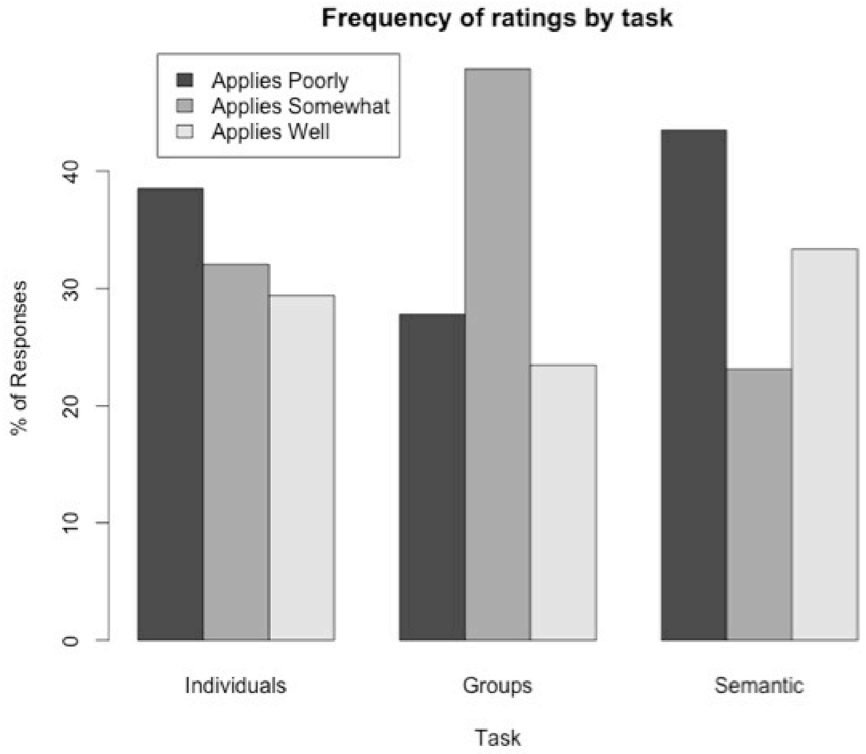


Figure 28. Behavioral responses for the three tasks. Bar heights represent percentage of responses within each task.

Supplementary Tables

Supplementary Table 1. White matter and ventricular activations from other tables. Because these are difficult to interpret, they are presented here, together, for clarity. Coordinates are peaks for contrasts, centers of mass for conjunctions.

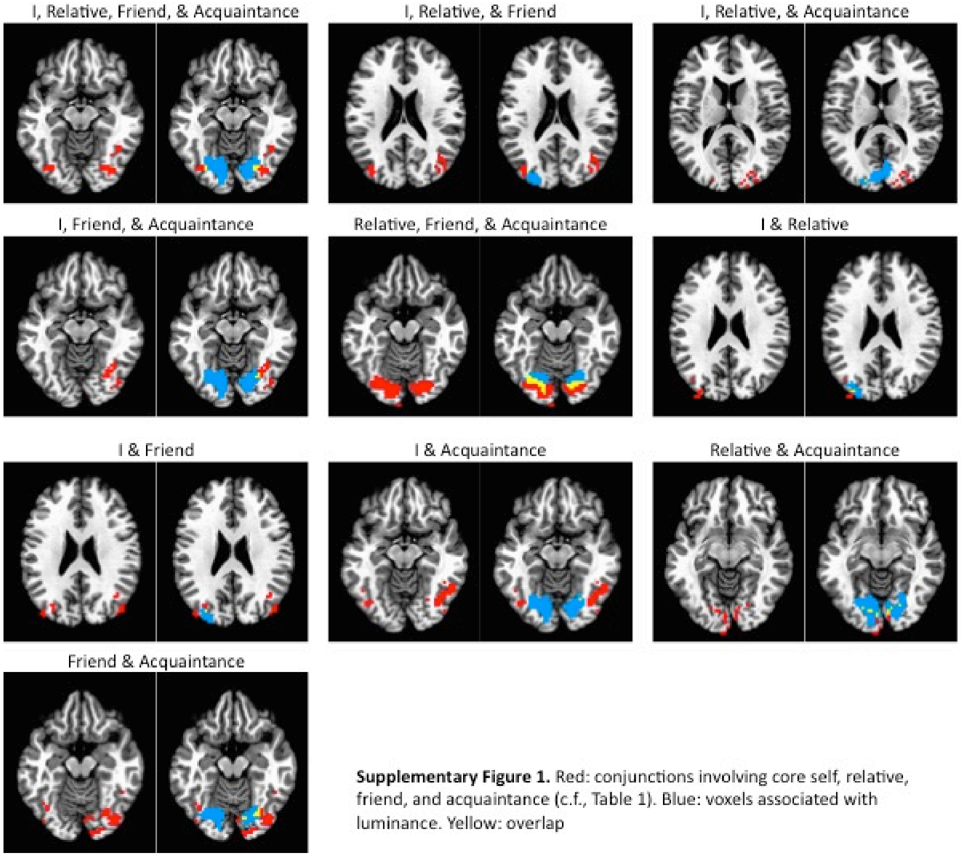
Lobe	**Medium**	**Table**	**Analysis**	**Volume**	**Peak t**	**(x, y, z)**
R frontal	white matter	1	I_1	**83**	NA	(23, -1, 28)
R frontal	white matter	1	I_1	**43**	NA	(17, 26, 16)
L frontal	white matter	1	I_1	**62**	NA	(-18, 33, 7)
L frontal	white matter	1	Relative_1	**43**	NA	(-24, -29, 28)
B lateral ventricle	CSF	1	I_I	**47**	NA	(3, -12, 20)
L lateral ventricle	CSF	1	Friend_I	**43**	NA	(-30, -41, 4)
L lateral ventricle	CSF	1	Acquaintance_I	**51**	NA	(-18, -12, 29)
L fronto-parietal	white matter	4	$\text{Individuals}_1 * \text{Groups}_1$	**83**	NA	(-21, -18, 30)
R fronto-parietal	white matter	4	Groups_1	**67**	NA	(20, -9, 28)
L lateral ventricle	CSF	4	Individuals_1	**103**	NA	(-21, -29, 21)
L fronto-parietal	CSF	4	Groups_1	**156**	NA	(-20, -24, 31)
B lateral ventricle	CSF	4	Prompts_1	**61**	NA	(4, -9, 22)
L frontal	white matter	7	$\text{Relative}_2 > I_2$	**47**	-5.36	(-28, -4, 30)

Supplementary Table 2. IOS ratings given by each participant for each of the 4 individuals (I, Relative, Friend, and Acquaintance). Because the IOS is a 1-7 closeness scale designed for other, non-self persons, the self was given a default rating of 8. These values were used to generate the amplitude-modulated regressors for closeness. The results of this analysis are presented in Table 3.

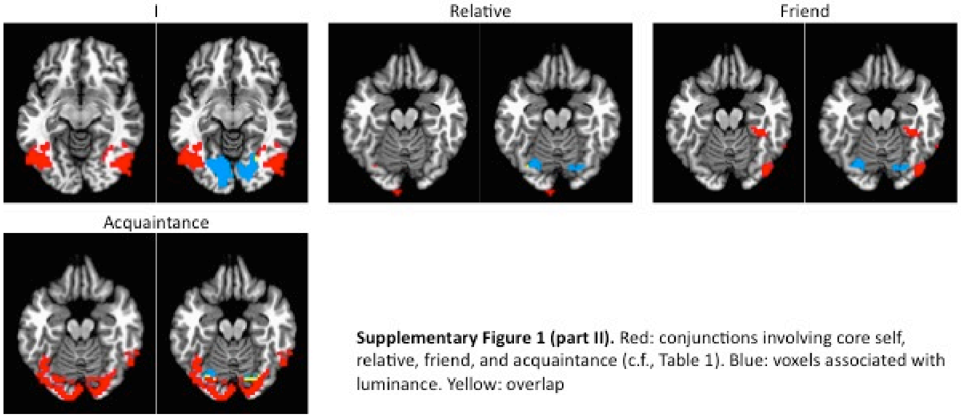
Participant #	**I**	**Relative**	**Friend**	**Acquaintance**
1	8	5	6	1
2	8	6	5	4
3	8	7	5	3
4	8	6	4	2
5	8	3	6	2
6	8	3	4	2
7	8	5	3	1
8	8	4	7	4
9	8	2	5	4
10	8	5	6	2
11	8	3	5	1
12	8	4	5	2
13	8	5	4	2
14	8	4	6	3
*15	8	6	4	2
16	8	6	5	2
17	8	6	5	3
18	8	7	5	2
19	8	5	4	2
20	8	7	5	2

***Note:** Participant #15 was removed form all analyses due to an incomplete session resulting from discomfort in the scanner.

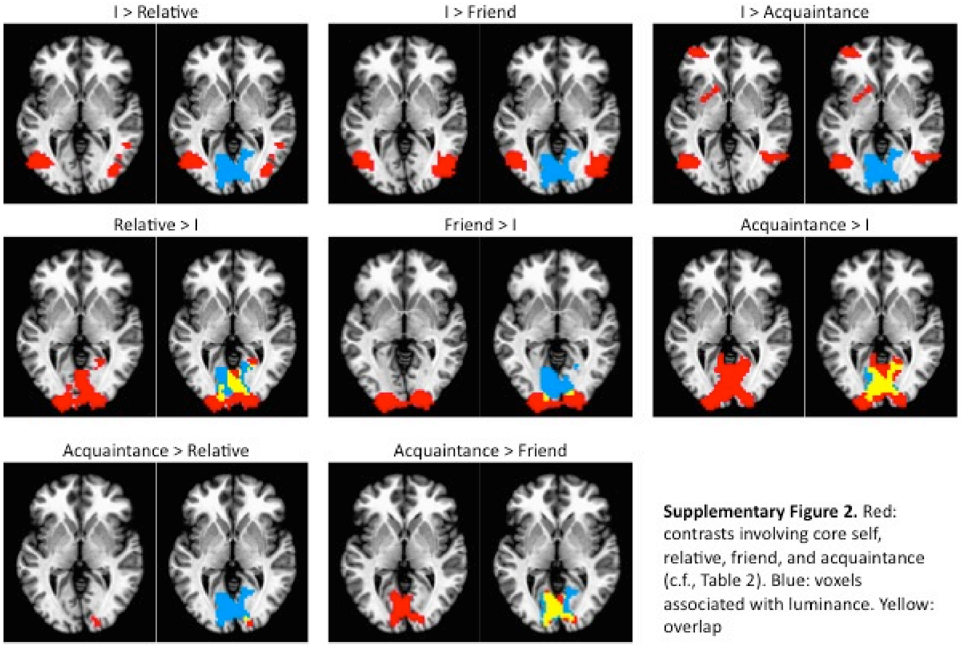
Supplementary Figures



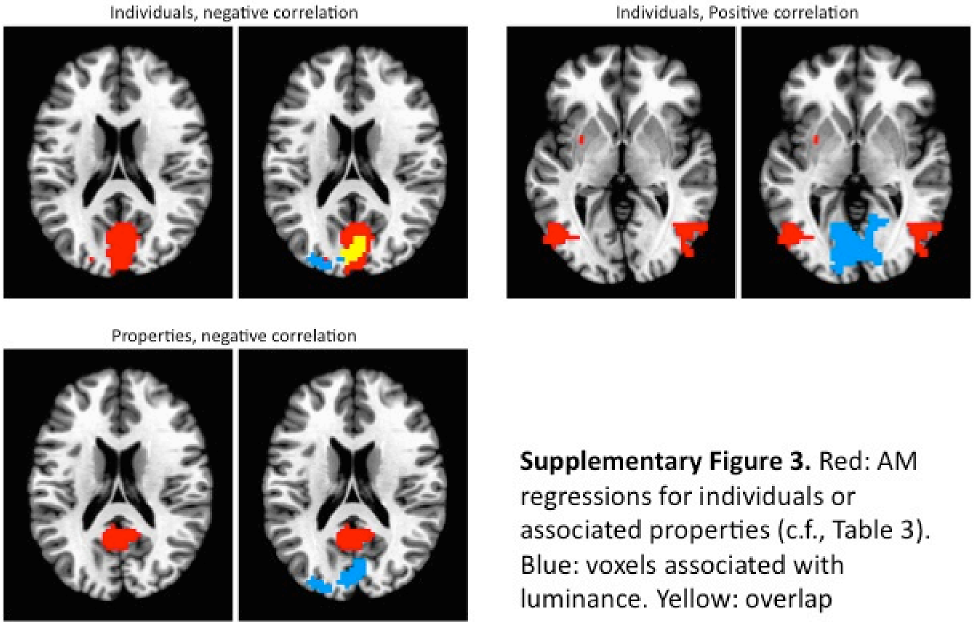
Supplementary Figure 1. Red: conjunctions involving core self, relative, friend, and acquaintance (c.f., Table 1). Blue: voxels associated with luminance. Yellow: overlap



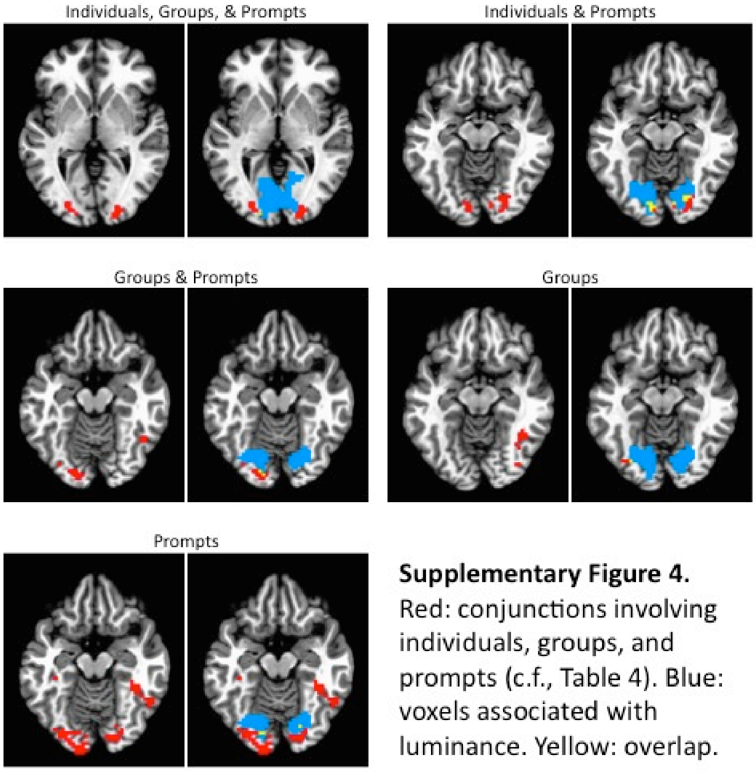
Supplementary Figure 1 (part II). Red: conjunctions involving core self, relative, friend, and acquaintance (c.f., Table 1). Blue: voxels associated with luminance. Yellow: overlap



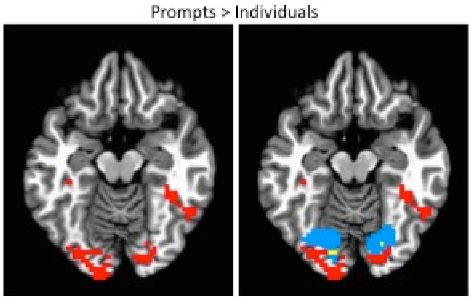
Supplementary Figure 2. Red: contrasts involving core self, relative, friend, and acquaintance (c.f., Table 2). Blue: voxels associated with luminance. Yellow: overlap



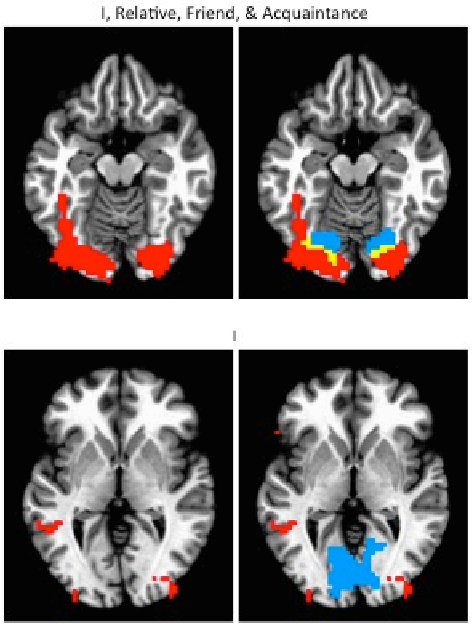
Supplementary Figure 3. Red: AM regressions for individuals or associated properties (c.f., Table 3). Blue: voxels associated with luminance. Yellow: overlap



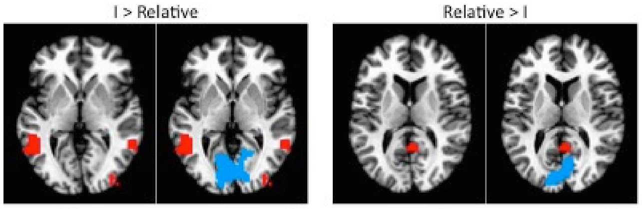
Supplementary Figure 4. Red: conjunctions involving individuals, groups, and prompts (c.f., Table 4). Blue: voxels associated with luminance. Yellow: overlap.


Prompts > Individuals

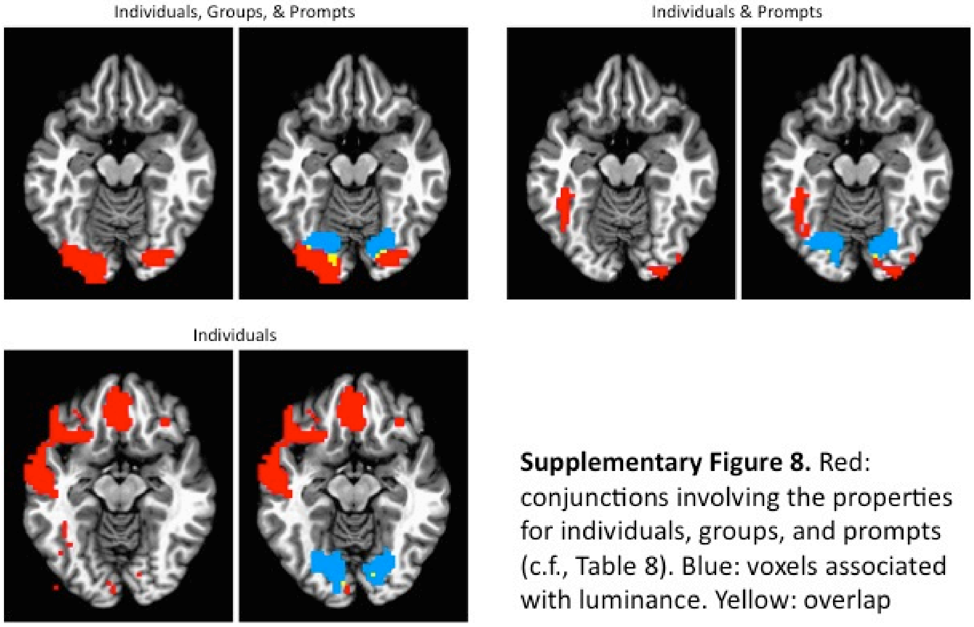
Supplementary Figure 5. Red: contrasts involving individuals, groups, and prompts (c.f., Table 5). Blue: voxels associated with luminance. Yellow: overlap



Supplementary Figure 6. Red: conjunctions involving the properties for I, Relative, Friend, and Acquaintance (c.f., Table 6). Blue: voxels associated with luminance. Yellow: overlap



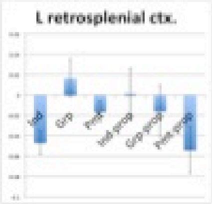
Supplementary Figure 7. Red: contrasts involving properties for I, Relative, Friend, & Acquaintance (c.f., Table 7). Blue: voxels associated with luminance. Yellow: overlap



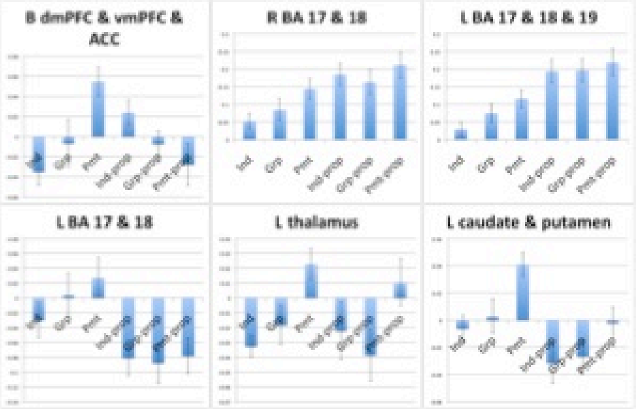
Supplementary Figure 8. Red: conjunctions involving the properties for individuals, groups, and prompts (c.f., Table 8). Blue: voxels associated with luminance. Yellow: overlap

9a. Individuals$_1$ > Groups$_1$

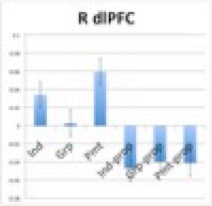
9b. Groups$_1$ > Individuals$_1$



9c. Prompts$_1$ > Individuals$_1$

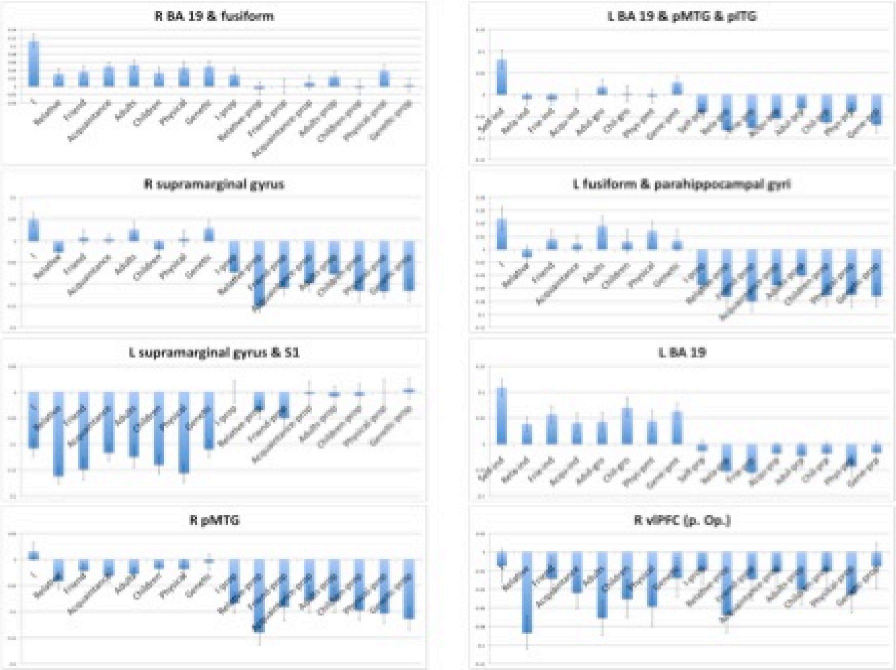


9d. Prompts$_1$ > Groups$_1$

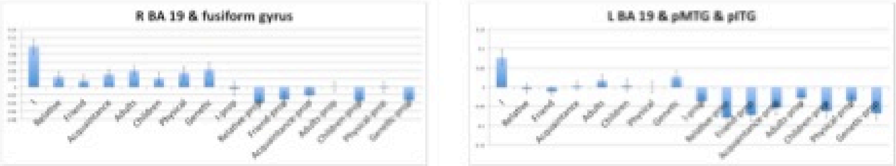


Supplementary Figure 9. Betas from each collapsed condition, in clusters from the Phase 1 contrasts *Individuals$_1$ > Groups$_1$ > Prompts$_1$*

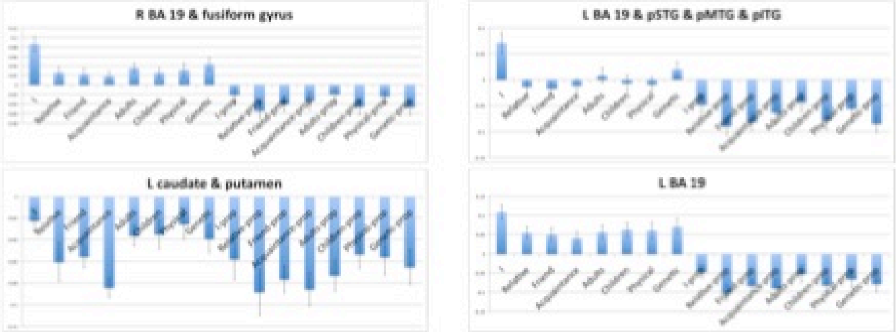
10a. $\mathrm{I}_1 > \mathrm{Relative}_1$

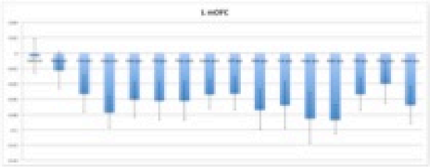


10b. $\mathrm{I}_1 > \mathrm{Friend}_1$



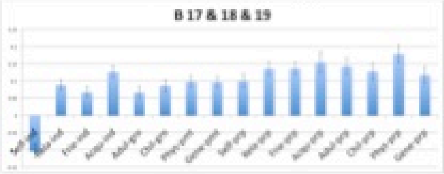
10c. $\mathrm{I}_1 > \mathrm{Acquaintance}_1$



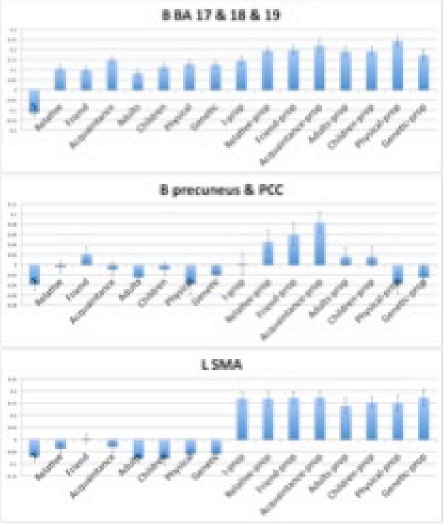


Supplementary Figure 10 (part I). Betas from each condition, in clusters from the Phase 1 contrasts between conditions for each individual: $I_1 > Relative_1 > Friend_1 > Acquaintance_1$.

10d. Relative$_1$ > I$_1$

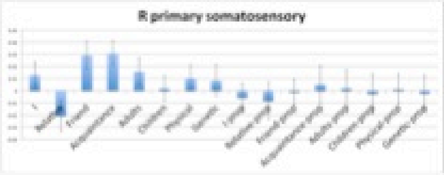


10e. Friend$_1$ > I$_1$

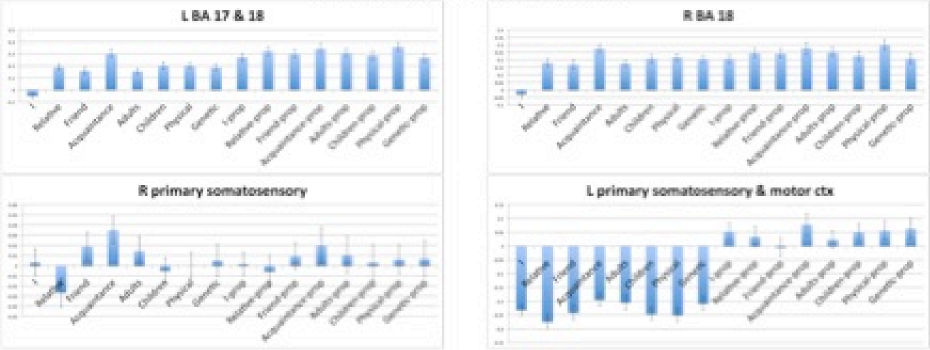


10f. Acquaintance$_1$ > I$_1$

10g. Friend$_1$ > Relative$_1$



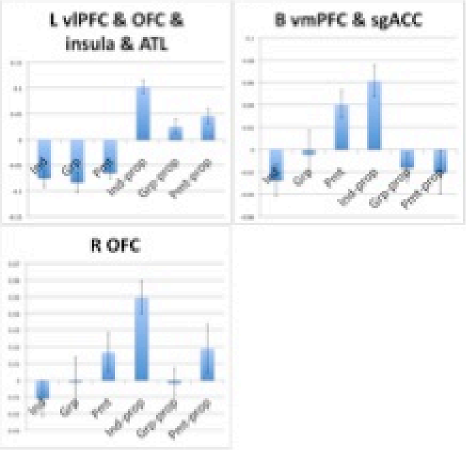
10h. Acquaintance$_1$ > Relative$_1$



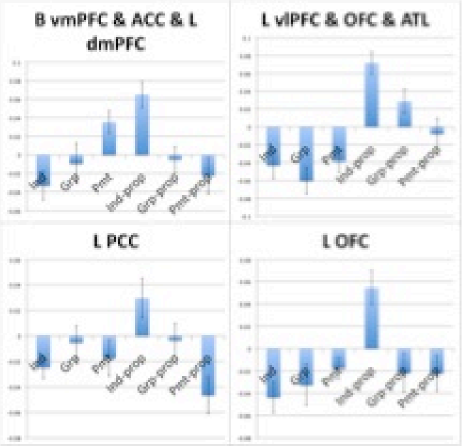
10i. Acquaintance$_1$ > Friend$_1$

Supplementary Figure 10 (part II). Betas from each condition, in clusters from the Phase 1 contrasts between conditions for each individual: *Acquaintance*$_1$ > *Friend*$_1$ > *Relative*$_1$ > *I*$_1$.

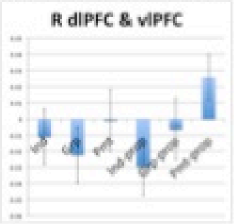
9a. Individuals$_2$ > Groups$_2$



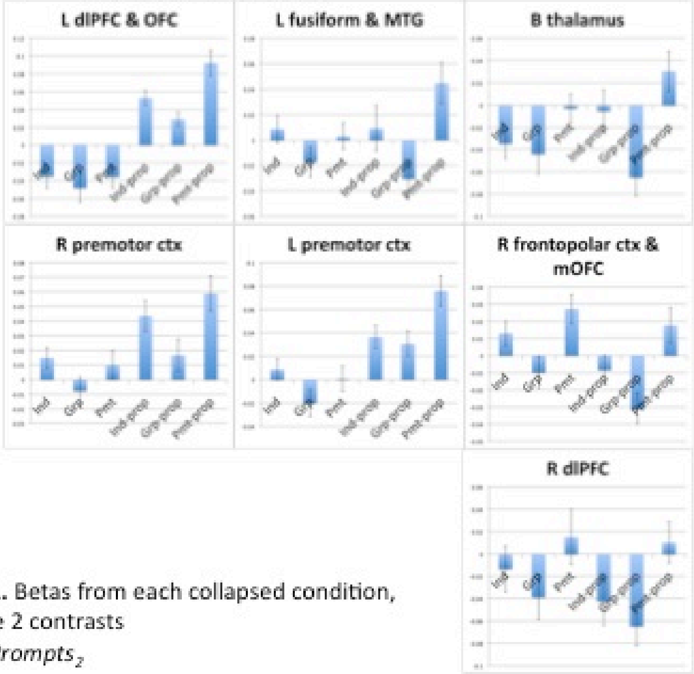
9b. Individuals$_2$ > Prompts$_2$



9c. Prompts$_2$ > Individuals$_2$

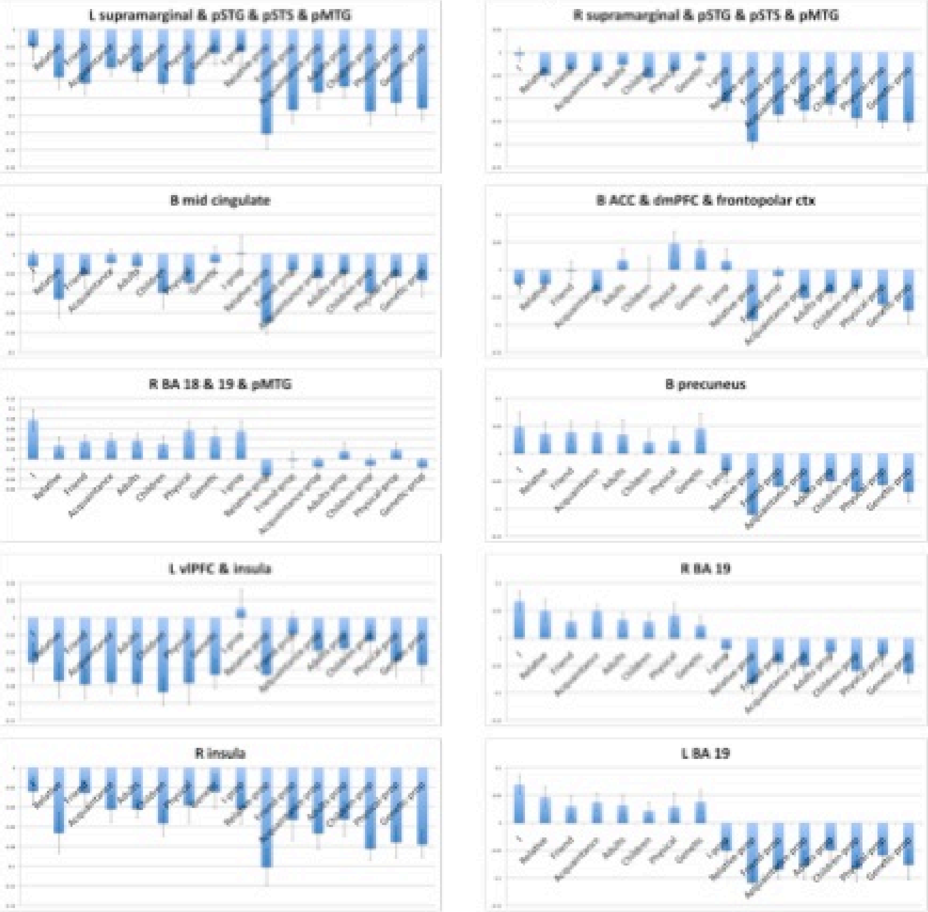


9d. Prompts$_2$ > Groups$_2$

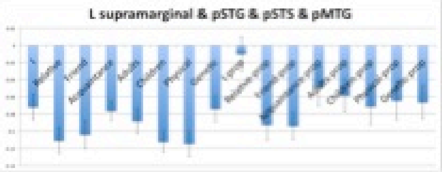


Supplementary Figure 11. Betas from each collapsed condition, in clusters from the Phase 2 contrasts *Individuals$_2$ > Groups$_2$ > Prompts$_2$*

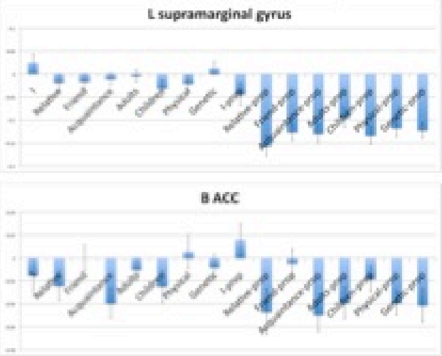
12a. I_2 > $Relative_2$



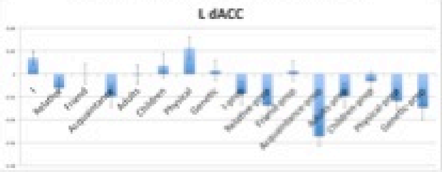
12b. I_2 > $Friend_2$



12c. I_2 > $Acquaintance_2$

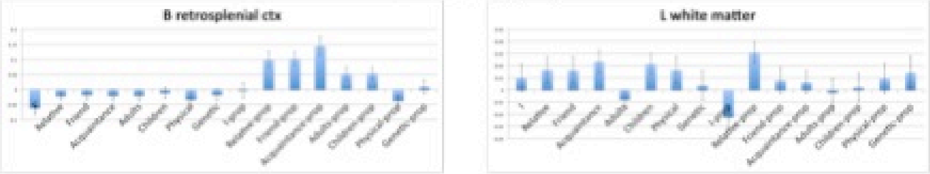


12d. $Friend_2$ > $Acquaintance_2$



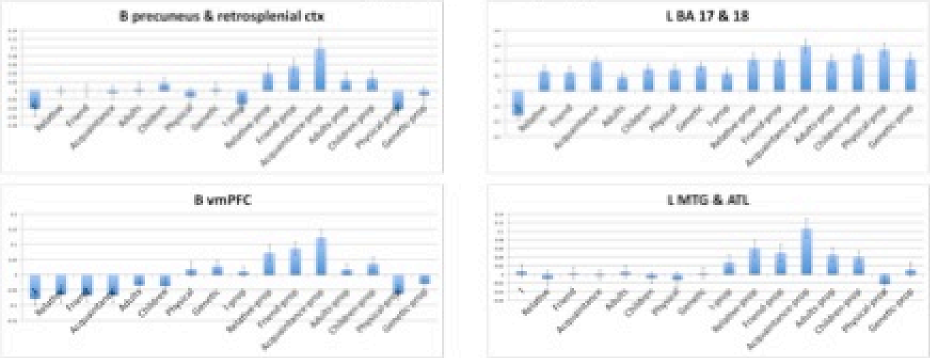
Supplementary Figure 12 (part I). Betas from each condition, in clusters from the Phase 2 contrasts between property conditions for each individual: I_2 > $Relative_2$ > $Friend_2$ > $Acquaintance_2$.

12e. Relative$_2$ > I$_2$

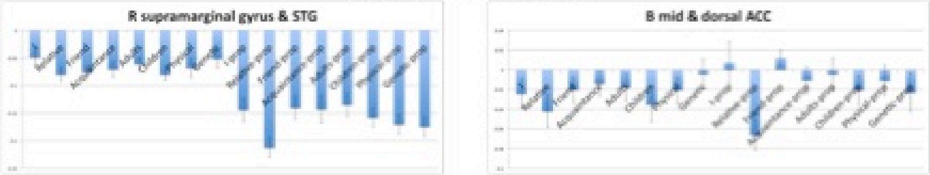


12f. Friend$_2$ > I$_2$

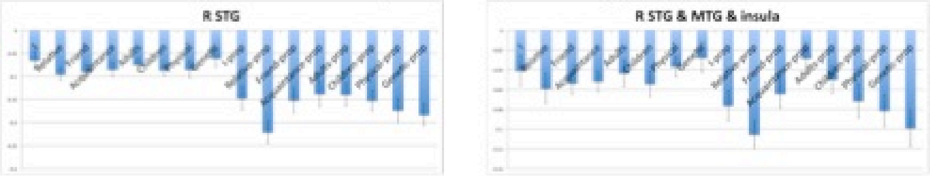
12g. Acquaintance$_2$ > I$_2$

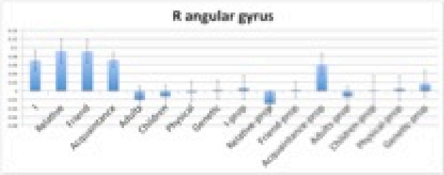


12h. Friend$_2$ > Relative$_2$



12i. Acquaintance$_2$ > Relative$_2$





Supplementary Figure 12 (part II). Betas from each condition, in clusters from the Phase 2 contrasts between property conditions for each individual: *Acquaintance$_2$ > Friend$_2$ > Relative$_2$ > I$_2$*.

Appendix 1: *Property* stimuli

96 properties were chosen for the norming study: 16 in each subclass (Figure 12). Of these, 36 were chosen for the main experiment according to the results of the norming study. These are in bold below.

adventurous
altruism
always on time
ambidextrous
ambition
artistic
big ears
brave
careful
caring
competence
confidence
considerate
cooperative
coordinated
creativity
curiosity
curly hair
curvaceous
dancer
dark-skinned
delicate
efficient
ethical
expressive
fast typist
firm handshake
flexible
forgiving
freckles
frequent traveler
friendliness
generosity
giggly
good mood
good posture
graceful
groaning
hairy
hard kicker
hardworking
high cheekbones
honesty
humble
inquisitive nature
knowledgeable
knuckle-cracking
large nose
liberal
light-haired
long arms
long-legged
modesty
muscular
nasal-sounding
nimble
non-judgmental
olive-skinned
open mind
organized
perfectionist
perseverence
quick runner
raspy
rational
responsibility
round face
self-discipline
selflessness
sense of duty
sentimental
short reach
silent
sincerity
sneezy
snoring
snorting
softspoken
spontaneity
square jaw
stomach growling
strong grasp
talkative
tall
thin
throat-clearing
ticklish
tolerance
trusting
trustworthy
unconventional
uninhibited
vivid imagination
wheezy
whispering
whistling

Appendix 2: Stimulus Generation - Norming Study and Analysis

Initial Stimulus Set

As a starting point, I copied over all of the human properties I had generated for the previous experiment (Drucker et al., submitted). This included not only the 60 properties I actually used, but also surplus properties that were excluded from that study. The concrete/abstract distinction was recast as a physical/mental distinction because I felt these terms were A) more transparent, and B) more in line with the neuroimaging literature upon which I am building (e.g., Mitchell, Banaji, & Macrae, 2005; Jenkins & Mitchell, 2011). I decided which properties fit which class (*mental* or *physical*) based on my own opinions, which would be refined by the norming study. I made no attempt to judge how "genetic" each property was. Some part-of-speech changes were made for counterbalancing purposes.

In all, I ended up with 96 counterbalanced properties. Half (48) were physical, and half were mental. Half were nouns, and half were adjectives. Equal numbers of properties belonged to each subclass. For mental properties, these subclasses were *conscientiousness*, *agreeableness*, and *openness to experience*. For physical properties, these subclasses were *visual*, *auditory*, and *sensorimotor*.

Norming

Purpose. In the semantic task of the main experiment, two questions were asked: "To what extent is the following property *physical*?" and "To what extent is the following property *genetic*?" In order to maximize the homogeneity (across participants, for the same items) and diversity (across items) of participants' responses, I conducted a norming study to collect participants' ratings of each of the 96 properties along these two (not necessarily independent) dimensions.

Design. The norming study consisted of two blocks. In each block, participants rated either how *physical* or *genetic* they found each property on a 7-point scale. The order of the two blocks alternated across participants, of which I recruited 24. In each block, all 96 properties were presented in random order. This paradigm was designed and implemented in MATLAB and R, and the resulting scripts were run in E-Prime 2.

Analysis. Properties were chosen so as to optimize the following criteria, in decreasing order of importance:

- Counterbalancing for part of speech, number of characters, and number of words was preserved.
- Sub-classes (Figure 12) were all represented in the final set.
- Distribution of *physicalness* and *geneticness* ratings were as uniform as possible.

All analyses, and the construction of lists and scripts for E-Prime, were carried out using Microsoft Excel spreadsheets and a suite of scripts designed for this experiment in MATLAB and R.

Designing the Main Experiment

The main experiment consisted of six functional runs, each containing four blocks (two individuals blocks, one groups block, and one semantic block: see Figure 9). Within runs, the blocks were ordered in every possible way that does not allow consecutive Individuals blocks. This rule allowed six possible orderings of blocks within a run, hence six runs.

The norming study selected 36 unique properties, with each of the following distinctions crossed with one another. Half were *physical* (rated above median on the *physicalness* question), whereas half were *mental* (rated below median). Half were *genetic* (rated above median on the *geneticness* question), whereas half were *learned* (rated below median). Each of these two p-classes (*physical* and *mental*) consisted of three subclasses as described earlier and in the main text. Because the subclasses were not exactly evenly represented, extra properties were assigned to the next best fit. So far, 12 degrees of freedom have been described (2 * 2 * 3 = 12). Each unique combination of p-class, g-class, and subclass defines what I call a *cohort*. The three remaining degrees of freedom (36 ÷ 12 = 3) were arbitrarily indexed by a value called *position*. Therefore, each cohort contained three properties: one at each position.

There were 36 unique properties, but 48 complete trials in a run. Therefore, 12 properties must be repeated in each run. If a property was repeated in a particular run, I enforced that it appeared exactly once in the Properties block and once elsewhere. Each property is repeated in exactly two different runs, with co-occurrences minimized (i.e., if *tall* and *adventurous* are both repeated in run 1, pains were taken to ensure that *tall* and *adventurous* were not repeated in the same run again. Of course, this were not possible in all cases, but I believe my solution was optimal).

Each block featured twelve properties: one from each cohort. Of those twelve, there were four from each of the three positions. There were some dependencies here (e.g., some blocks are all position 1), but that does not matter because position is arbitrary. At this point, the order within each block was randomized and the particular properties generated in the norming study were assigned to their appropriate places in each block.