The Conceptual Content of Mental Activity

Jeffrey R. Binder

This chapter discusses some phenomenological and biological links between mentalizing and general concept retrieval. As attested by this book, the neural underpinnings of our ability to hypothesize about the mental content of other intentional beings has become a topic of great interest in psychology and neuroscience. The central importance of this ability in everyday human life reflects the myriad survival advantages it conveys, which are likely reflected in somewhat specialized neurobiological representations. I argue, however, that the neural systems supporting these representations also support other types of conceptual content, placing a substantial burden of proof on any claims for functional specialization.

The ability to store and use knowledge about the world is a core feature of the human brain that has been central to our evolution and survival success, making it possible to reliably avoid known dangers and anticipate future needs by planning. People have spread across the globe and flourished through the invention of technology, including such seminal inventions as constructed shelters, farming, domestication of animals, methods for storing and preserving food, and devices for capturing and transforming energy. In each of these cases, known facts about objects and observed events were mentally manipulated, analyzed, and synthesized to create novel methods for enhancing survival. Today most adults use the same processes on a daily basis to make short- and long-term plans for beneficial future activities and solve small-scale problems. Creative analysis and synthesis of stored knowledge is used on a daily basis in the social sphere to resolve conflicts, communicate ideas, and organize groups of people.

In addition to providing a means of meeting the various exigencies of daily life, activation and manipulation of stored conceptual information provides a mechanism for such important mental activities as pleasurable recall, daydreaming,

---

J. R. Binder (✉)
Department of Neurology and Biophysics, Medical College of Wisconsin,
Milwaukee, WI, USA
e-mail: jbinder@mcw.edu

M. Gilead, K. N. Ochsner (eds.), *The Neural Basis of Mentalizing*,
https://doi.org/10.1007/978-3-030-51890-5_20

reflection on art and culture, and analysis of one's own behavior and emotional responses.

Concepts and the Content of Mental Experiences

I argue that the primary contents, or "intentional objects" (Husserl, 1973/1900), of mental experience are concepts. A concept is a mental representation (which may be relatively simple or complex) resulting from generalization over many similar experiences, capturing what is common to these experiences. The concept of a concrete object like *dog*, for example, is an idealized or schematic representation of the characteristics of previously experienced dogs. Concepts like *dog* are referred to as category-level concepts because they refer to a set of unique individuals. Representations of particular individuals (e.g., *my dog Luna*), however, are also generalizations from experience and therefore concepts. Concepts have defining intrinsic features (e.g., shapes, colors, parts, movements, sounds), but also exist within a complex network of other associated concepts. The concept *dog*, for example, may have associations with concepts like *friend*, *love*, *loyalty*, *leash*, *bone*, *walk*, *breed*, *pedigree*, etc. established through co-occurrences in complex verbal and nonverbal experiences.

Concrete object concepts with verbal labels, like *dog*, have dominated much of the theoretical and empirical work on concepts (particularly in the neuroimaging world), but our vast store of concepts also includes concrete entities that are not objects (*air*, *water*, *soil*); concrete actions and events (represented in language mainly by verbs and sentences, but also by nouns like *party* and *explosion*); entities occurring as mental experiences (emotions and thoughts); quantity concepts (number, duration, and size); complex social/behavioral constructs (*honor*, *loyalty*, *democracy*, *justice*); cognitive and scientific domains (*geometry*, *law*, *philosophy*); spatial, temporal, and causal relation concepts; and many other categories. Because not all experiences are labeled with words, not all concepts have a name. The experience of satisfaction from another person's misfortune, for example, is an unnamed concept for English speakers who have not learned the word *schadenfreude*. In particular, many perceptual categories, acquired from generalization over repeated experiences, exist for which we have no names (e.g., the characteristic head shape of a particular kind of animal).

Activating a concept in the mind involves neural processing in a widely distributed brain network that represents (i.e., stores in long-term memory) and retrieves conceptual knowledge (Binder, Desai, Conant, & Graves, 2009). Since the mid-century "cognitive revolution," concept representations in the brain have been portrayed as highly abstract and localist, much like symbols in a computer program (Pylyshyn, 1984), and many authors still advocate at least a partial role for abstract representations in conceptual cognition (Dove, 2009; Mahon & Caramazza, 2008). Much behavioral and neuroimaging evidence suggests, however, that activating a concrete concept also entails activating perceptual representations of the concept in various sensory-motor modalities, such as information about its visual, tactile, auditory, or associated action features (Fernandino et al., 2016; Kiefer & Pulvermüller, 2012; Meteyard, Rodriguez Cuadrado, Bahrami, & Vigliocco, 2012). The degree to which this perceptual information becomes activated and enters awareness appears to depend on task demands. At the extreme, a visual or other sensory image may appear in awareness, but such "imagery" phenomena are best understood as a manifestation of sustained concept activation rather than a qualitatively distinct process. On this view, information stored in the brain about modality-specific (i.e., visual, auditory, tactile, action) attributes of concrete objects and events is not somehow separate from the concept representation, rather it *is* (at least a large part of) the concept representation.

The central role of concept retrieval in communication is uncontroversial: What is the purpose of communication if not to transmit concepts? If an acquaintance says, for example, "We had a good tennis game last week," it is obvious that understanding this message requires retrieval of basic knowledge about the concepts *we*, *had*, *good*, *tennis*, *game*, *last*, and *week*, and about the more specific concepts *tennis game* and *last week*. From this information, you, the hearer, might construct a mental image of the tennis game you had with the speaker, and respond by communicating labels for concepts like *I* and *agree*. Activation of conceptual knowledge, however, is not confined to the domain of verbal communication. A long tradition in linguistics and psychology linking concepts with words has obscured the fact that concept retrieval is a ubiquitous and core feature of nearly all mental activity. The paragraphs that follow discuss this point in relation to several cognitive domains usually considered to be distinct from general concept retrieval processes, all of which show considerable overlap in neuroimaging studies with both general concept retrieval networks and mentalizing networks.

Retrieval of personal episodic memories is traditionally distinguished from retrieval of concepts (semantic memory), but I argue that episodic memories are composed almost entirely of concepts. Consider that retrieval of the detailed sensory-motor events that occurred during the aforementioned tennis game, even if that were possible, would not be sufficient in itself for episodic memory retrieval. A particular set of sensory-motor events can only be recognized *as* a tennis game by retrieving the concept *tennis game*. Put another way, "understanding" always involves concept retrieval, and concepts exist in the brain to provide understanding. In the case of episodic memories, what is mainly remembered are not the detailed sensory-motor events that occurred, but an abstract version of events composed of concepts with varying amounts of perceptual detail. Episodic memory might be more properly seen as a particular kind of knowledge manipulation that creates spatial-temporal configurations of concepts representing objects, events, and other entities, including cognitive and affective phenomena.

"Autobiographical" memory is even more clearly dependent on concept retrieval, for in this case the original events have been stripped of nearly all perceptual detail and are remembered mainly as facts (e.g., place of birth, childhood home, education history). To say, "I was born in Chicago" is not to claim any perceptual memory for the events of the birth, but rather to retrieve the concepts of *birth*, *in*, and *Chicago* and to self-identify with this combination of concepts, where *I* and *self* are also nothing more or less than concepts.

The notion of autobiographical memory retrieval has relevance to the notion that some mental experiences engage a "concept of self" (Gillihan & Farah, 2005; Vogeley et al., 2001). In addition to autobiographical facts, the self-concept includes knowledge about one's own beliefs and values, likes and dislikes, physical and cognitive characteristics, relationships to others, financial situation, personal goals, and so on. By definition, such information is of great personal relevance, and the ability to retain and retrieve such information seems to be a logical prerequisite for everyday decision-making. How could I plan my day-to-day activities without knowing my own preferences, abilities, and goals? Yet to claim self-referential processing as a special mental activity separate from concept retrieval seems difficult to justify. Are physical traits, cognitive abilities, values, relationships, and goals not concepts? To agree or disagree, for example, with the statement "I value financial independence" surely depends on the ability to retrieve a representation of the various concepts expressed in this proposition, and probably on retrieval of a wide range of associated concepts, like *parents* and *job*. As mentioned above, even the notion of "I, myself" is a concept, if an elemental one formed at a very early stage of cognitive development. The view that self-processing arises from association of the "I, myself" concept with other concepts unpacks and demystifies this seemingly special mental ability, revealing it to be yet another instance of concept retrieval and association.

Prospection, i.e., imagining the future, is often held to be a prominent component of mental experience (Ingvar, 1985; Schacter & Addis, 2007). As was the case with imagining past events (episodic memory retrieval) and reflecting on one's concept of self, it is difficult to see how imagining future events could proceed without the core process of concept retrieval. A useful example is the participant told to "rest" in an fMRI experiment, who uses this time to consider available options for dinner after the scanning session is finished. Given our essential status as animals who benefit from the ability to store, recall, and assess food sources, it seems likely that this particular example of "future planning" has extensively evolved over the eons and provided important survival advantages. Even a cursory consideration of the processing involved, however, reveals this seemingly "special" activity of prospection to be little more than activation and evaluation of a set of related concepts. The varieties of possible cuisine, the specific shops or restaurants available and their pros and cons, the time available for a meal, the specific companions one expects to dine with and their preferences, relative differences in cost—all of these are concepts formed by generalizations from prior experiences. As with episodic memory retrieval and self-oriented cognition, imagining future scenarios cannot logically be separate from retrieval of the concepts that comprise the actual content of these mental experiences.

Working with Concepts: Selection, Analysis, and Synthesis

As outlined briefly above, mental activities generally involve the retrieval, or reactivation, of concept representations. For most such activities, however, the brain processes involved go beyond mere concept activation and include selection and manipulation of activated concepts. Selection refers to the enhanced activation, probably through an attentional mechanism, of a concept or concepts that are of greatest relevance and usefulness in a given circumstance, from among a larger set of activated concepts (Badre, Poldrack, Pare-Blagoev, Insler, & Wagner, 2005; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). In a naming task, for example, a picture of a sheep might activate a field of concepts like *sheep*, *lamb*, *goat*, *cow*, etc., requiring selection of the most appropriate response from among these competitors. Although studied almost exclusively in the setting of overt tasks, concept selection is a basic component of all conceptual processing and likely occurs even during "spontaneous" mental activity. Consider the fMRI participant planning dinner during a "resting" interval in the scanner: concept selection occurs at every phase of this mental activity, from the focus of attention on *dinner* as opposed to other meals, to selection of *restaurants* as the search domain as opposed to other types of establishments, to the use of certain criteria and not others for assessing restaurant options, to selection of some people and not others as potential companions, and so on. This classic prospection task might be redefined (somewhat arbitrarily) as a "self-processing" task if the participant plans to dine alone and therefore focuses exclusively on self-preferences. In addition to selection of concepts like *dinner* and *restaurant*, the focus on self requires selection of "self vs. other" preferences and self-preference criteria to be given the most weight. Selection mechanisms likely also play a role during recall of personal episodic memories. Such memories are not holistic, indivisible entities, but are made up of spatiotemporal configurations of object and event concepts. The experience of such a memory typically leads to attentional focus on certain aspects of the memory and not others, i.e., selection, which determines the course of subsequent episodic recall or prospective thinking.

Analysis refers to the delineation of component features of concepts. Concepts are nearly always composed of simpler elements, such as parts of objects, distinguishable sensory features of objects, separable parts of actions, participants in events, and sequential steps within events. To solve a problem or formulate a plan, it is often necessary to decompose a retrieved concept into its component parts. Deciding which car to purchase from among many options, for example, requires analysis of the concept *car* into components like *shape*, *color*, *size*, *mileage*, *reliability*, *safety*, etc. Planning a birthday party requires analysis of the concept *birthday party* into components like *invitees*, *invitations*, *location*, *cake*, *candles*, *presents*, and so on. Each of these components is also a concept, therefore what appears intuitively to be an analytical or "breaking apart" process might be better understood as a process of activating a field of associated concepts that stand in a part-whole relationship to the parent concept. In some ways, this process is opposite to selection: whereas selection aims to focus attention on a single concept by suppressing activation of related concepts, analysis aims to activate a field of closely related concepts.

Synthesis refers to the construction of new concepts, including plans for future actions, by assembling components within schemas. A schema is a representational framework that organizes category types and relationships (Rumelhart, 1980).

Schemas are used for mental organization of complex concepts, such as events involving social interactions and spatial-temporal sequences, as well as simple object concepts. A concept like *fruit*, for example, can be represented by a schema composed of "slots" for shape, size, color, taste, juiciness, seed-type, etc. A complex concept like *party* might employ a schema with slots for location, purpose, time and duration, types of attendees, sub-events during the party and their order of occurrence, etc. We use schemas to organize and understand everyday experiences by fitting features of those experiences into pre-learned schema, sometimes leading to prejudice, confirmation bias, and other effects of stereotypical thinking (Bartlett, 1932). We use schemas to plan simple and complex behaviors, typically with slots for goal(s), actor(s), instrument(s), action(s), and patient(s) (Minsky, 1975).

Mentalizing as a Conceptual Activity

Hypothesizing about the content of other people's thoughts and motivations is arguably a special case of the more general processes of concept retrieval, concept selection, analysis, and synthesis. From infancy we discover that we have needs that must be filled, like hunger, thirst, affection, physical comfort, sleep, and safety. We also discover various means of meeting these needs, and because these needs and means of fulfillment recur many times in many situations, they become generalized concepts that we use, consciously or not, to formulate actions. The toddler's statement "Mommy I'm hungry" is a demonstration that the child has learned the concepts *I*, *hungry*, and *mother* and is able to select these concepts from among a field of related ones like *you*, *thirsty*, and *brother*. Analysis is demonstrated by the child's knowledge that, along with her other characteristics, *mother* is a *giver* of *food*. Through multiple experiences in which mother (or someone else) provides the child with food, a schema develops in which the child expresses (verbally or nonverbally) a need to someone, who responds by providing something to meet the need. Synthesis occurs when the concepts *hungry* and *mother* are fit into this general schema, creating an action plan.

By the time a child is able to formulate such a plan, another critical concept will likely have been learned: the concept of having a mental plan. Concepts are generalizations learned from repeated experiences. I argue that any animal who can repeatedly form mental plans and experience the state of holding in mind a mental plan will eventually develop a concept of what it is to have a mental plan. The "experience" component is critical here. Artificial intelligence devices can be programmed to formulate action plans, though our intuition tells us that this ability alone doesn't create in the device an "experience" of having formulated a plan. Human (and many other animal) brains are different in this critical regard: we automatically extract from complex neural activation patterns a simplified representation that can be held in short-term memory and presented to "awareness." But this general abstraction process is the same whether the raw neural activation pattern results from an external sensory stimulus, an emotional response, or a mental event.

Learning the concept *I want* or *I believe* is not essentially different, in neurobiological terms, from learning the concept *red* or *heavy*.

Now consider what the toddler who says “Mommy I’m hungry” knows about his mother’s mental contents. It is quite likely that these contents are complex, probably including thoughts about other things she needs to do, how much food there is in the house, why her toddler is hungry so often, how much fun she had last night with her friends, etc. The toddler, on the other hand, knows only that mommy intends to get him food. How does he know this? Because he has learned from his own mental experiences the concept of having a mental plan, and he has observed on many occasions his mother executing the action of bringing food. Though not articulated overtly, the child knows (or at least expects) that his mother intends to bring him food once his own action plan (“Mommy I’m hungry”) has been executed. The fact that the child has no knowledge of the many other contents of his mother’s mind is proof that such contents must be learned through generalization over many similar experiences.

These general principles extend to all concepts acquired in the domain of social and emotional cognition. As we experience our own mental states, whether these involve desires for basic needs, emotional responses, thoughts, or simply curiosity about the environment, these recurring mental experiences evolve into generalized concepts that can be identified and articulated. Included among the core components of these concepts are our own responses and actions that result from these internal states, such as facial and body gestures that reflect emotional responses, actions taken to fulfill needs, and verbal expressions (words and phrases) that communicate the contents of our mental experience. Once conceptualized, these associated responses can be recognized in others, allowing us to infer the mental states that led to the responses, providing the basis for theory of mind. In addition to inference based on observation of others’ overt responses, we identify through experience the reliable environmental contexts that give rise to particular mental states, which then become associated with those states and can be used as additional evidence to infer mental states in others. A child’s own experiences with the emotional response caused by having a treasured toy taken away, for example, produces an association between this environmental context and the emotion of anger. A simple schema develops in which a negative emotion is experienced by sudden loss of an object. Observing another child in the same situation allows a kind of pattern completion to occur in the observer, in which the observed loss activates this previously learned schema and a representation in the observer of the likely emotional response that will occur in the other child.

The main point is that complex mental and behavioral phenomena reflecting the fact that we can infer the mental content of other intentional beings are the result of nothing more than learning through generalization over repeated similar experiences. I have elsewhere addressed the possible experiential origins of the concept of “animacy” (more properly, intentionality), which is critical for limiting the domain of possible entities to which theory of mind schema can be applied (Binder et al., 2016). We do not attribute mental states and intentions to inanimate objects, for example, because these objects do not move, show emotional responses, or communicate like intentional beings. Like our knowledge of mental states, action categories, and response schema, our knowledge of intentionality is a conceptual representation that can be activated, selected, analyzed, and synthesized with other concepts to produce action plans and inferences. Claims about processing in the domain of mentalizing and social cognition should recognize the essentially conceptual nature of these behaviors and the possibility that they are particular examples of computations (complex though they may be) arising within a more general conceptual system.

Neuroimaging Considerations

Functional neuroimaging evidence on brain systems supporting mentalizing have been expertly reviewed elsewhere (Mahy, Moses, & Pfeifer, 2014; Mar, 2011; Molenberghs, Johnson, Henry, & Mattingley, 2016; Van Overwalle, 2009) and by other contributors to this volume. Core nodes of this network include the "temporoparietal junction" (an ambiguous anatomical label usually referring to angular or supramarginal portions of the inferior parietal lobe), superior temporal sulcus, medial prefrontal cortex, lateral anterior temporal lobe, and posterior cingulate cortex. As pointed out by several authors (Andrews-Hanna, 2012; Buckner, Andrews-Hanna, & Schacter, 2008; Schilbach et al., 2012; Spreng, Mar, & Kim, 2009), this network overlaps extensively with the "default mode" network and with brain regions implicated in episodic and autobiographical memory retrieval, prospection, self-processing, and moral judgments. These latter overlaps lend support to proposals that memory retrieval, prospection, and self-processing are key components of the mental activity occurring during "resting" states (Andrews-Hanna, 2012; Buckner et al., 2008; Schacter & Addis, 2007). But what is the underlying reason for these overlaps, and why are mentalizing processes supported by virtually the same brain regions that support these other cognitive processes?

As discussed above, all of these mental activities depend on the core processes of activating stored concepts, concept selection, concept analysis, and schema-based synthesis. Sometimes ignored by social cognition researchers is a large parallel literature on single-word semantic processing showing that all of these brain regions are activated by simple contrasts like (word > matched pseudoword) and (conceptual task > matched phonologic task) (Binder et al., 2009) (Fig. 1). These contrasts, which typically use simple lexical or semantic decision tasks and neutral words drawn from a mix of conceptual categories, highlight domain-general brain areas involved in the basic processes of concept storage, retrieval and selection, analysis, and synthesis. The extensive overlap between these areas and those identified in mentalizing and other social cognition studies supports the idea that mentalizing, like most other mental activities, depends to a large extent on these domain-general conceptual processes.

A critical feature of this network is that it responds in proportion to the amount of conceptual content being processed (Binder, 2016). Activation in these areas

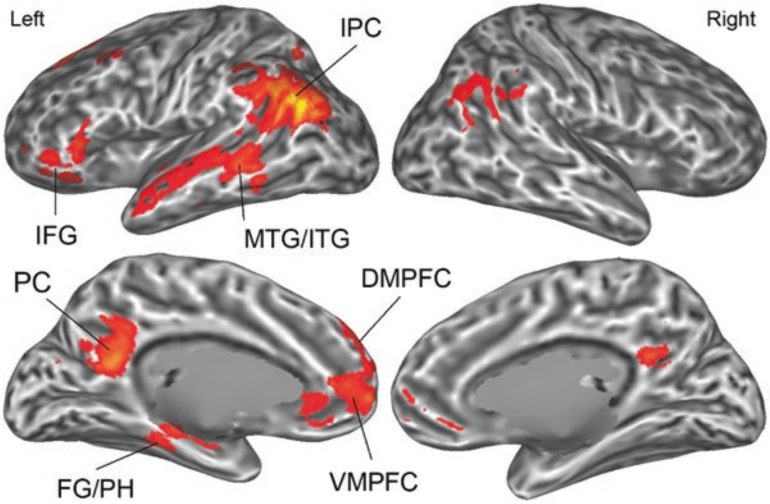

Fig. 1 A conceptual network identified by quantitative meta-analysis of 87 neuroimaging studies of semantic processing. The studies all included a manipulation of stimulus meaningfulness but no manipulation of modality-specific content. (Adapted with permission from Binder et al., 2009.) *DMPFC* dorsomedial prefrontal cortex, *FG/PH* fusiform gyrus/parahippocampus, *IFG* inferior frontal gyrus, *IPC* inferior parietal cortex, *PC* posterior cingulate/precuneus, *VMC* ventromedial prefrontal cortex

reflects the number of concepts that are active (and their intensity of activation) at any given moment, which in turn depends on the number and strength of associations that these concepts have. Distributed neural ensembles in these regions are literally equivalent to concept representations, each of which can activate a set of associated neural ensembles. All else being equal, a concept that activates many other associated concepts (causing, in turn, activation of the concepts associated with those concepts, and so on) will produce greater activation in these areas than a concept with relatively few or relatively weak associations (Bar, 2007). As mentioned above, nodes in this network are activated by single words relative to pseudowords (Binder et al., 2003; Binder, Medler, Desai, Conant, & Liebenthal, 2005; Henson, Price, Rugg, Turner, & Friston, 2002; Ischebeck et al., 2004; Kotz, Cappa, von Cramon, & Friederici, 2002; Kuchinke et al., 2005; Mechelli, Gorno-Tempini, & Price, 2003; Orfanidou, Marslen-Wilson, & Davis, 2006; Rissman, Eliassen, & Blumstein, 2003; Xiao et al., 2005). According to the present theory, this is due to the fact that pseudowords have no strong associations with concepts. Very similar results were obtained in studies comparing responses to familiar and unfamiliar proper names (Sugiura et al., 2006; Woodard et al., 2007). Like pseudowords relative to words, unfamiliar names, which refer to no known individual, have far fewer associations than familiar names, which refer to actual people about which one has associated knowledge.

Other observations explained by this general principle include activation of many of these regions by concrete relative to abstract concepts (Bedny &

Thompson-Schill, 2006; Binder et al., 2009; Binder, Medler, et al., 2005; Binder, Westbury, Possing, McKiernan, & Medler, 2005; Fliessbach, Wesi, Klaver, Elger, & Weber, 2006; Graves, Desai, Humphries, Seidenberg, & Binder, 2010; Jessen et al., 2000; Sabsevitz, Medler, Seidenberg, & Binder, 2005; Wallentin, Østergaarda, Lund, Østergaard, & Roepstorff, 2005) and frequently-used compared to infrequent words (Carreiras, Riba, Vergara, Heldmann, & Münte, 2009; Graves et al., 2010; Prabhakaran, Blumstein, Myers, Hutchison, & Britton, 2006). Concrete words show a variety of behavioral processing advantages over abstract words, including faster response times in lexical and semantic decision tasks and better recall in episodic memory tasks, reflecting the fact that concrete concepts more readily or automatically activate mental images and situational and contextual associations than abstract concepts (Paivio, 1986; Schwanenflugel, 1991). Word frequency is correlated with the number and strength of associations people generate in free association tasks (Nelson & McEvoy, 2000) and with the number of semantic features people produce in feature listing tasks (McRae, Cree, Seidenberg, & McNorgan, 2005). Assuming that words with higher frequency of use automatically activate a larger number of conceptual associations, frequency-dependent activation of the conceptual network is consistent with the aforementioned word-pseudoword, familiar-unfamiliar name, and concrete-abstract effects, all of which can be accounted for by a common underlying mechanism, i.e., relative differences in the overall intensity of activation of associated concepts.

These well-documented modulatory influences should be considered in interpreting functional imaging studies that aim to identify domain-specific processing. It is not hard to imagine, for example, the possibility that stimuli intended to specifically engage a theory of mind network might simply activate more or stronger conceptual associations than non-ToM stimuli, due to greater complexity, familiarity, or imageability, or to stronger engagement of attention by the ToM stimuli. There is no question that social interactions are an extremely important facet of our daily lives, and that we therefore know a great deal about and habitually pay close attention to human behavior. But this extended and readily accessible database of social knowledge creates an important potential confound in studies comparing processing of social vs. non-social stimuli. Are the activations observed in such comparisons specifically due to processing of social knowledge per se, or simply to stronger engagement of conceptual knowledge in general?

A concrete example of this type of confound can be found in experiments comparing verbal descriptions of complex social interactions (ToM stories) with vignettes lacking such interactions. In an item-level analysis, Dodell-Feder, Koster-Hale, Bedny, and Saxe (2011) noted substantial variation in magnitude of activation of the temporoparietal junction *within* the ToM and non-ToM conditions. That is, some ToM stimuli produced strong activation of the TPJ whereas others did not, and some non-ToM stimuli activated the region as strongly as or stronger than some ToM stimuli. As the authors noted, such variation suggests that other (non-hypothesized) stimulus features are modulating the activation. The authors considered 19 features, including 13 linguistic features (number of words per story, Flesch reading ease, anaphor reference, causal content, causal cohesion, lexical

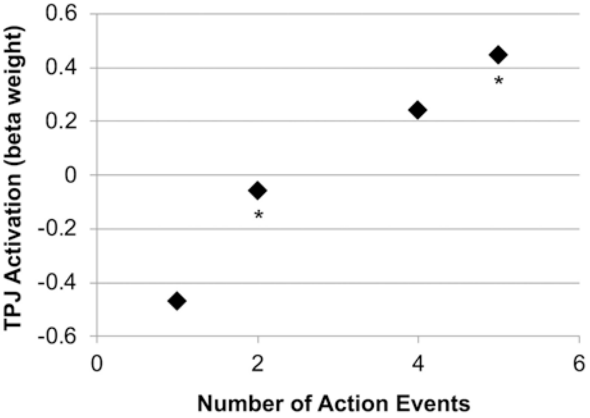

Fig. 2 Temporoparietal junction fMRI activation level produced by four story stimuli as a function of the number of action events described in each story. Items marked with an asterisk were theory-of-mind stories; unmarked items were stories describing physical events. The data are taken from examples provided in Dodell-Feder et al. (2011), Table 1

concreteness, negation, noun-phrase modification, higher-level constituency, number of words before the main verb, intentional content, attitude predication, and modality), 4 social features (number of people per story, the extent to which the items made readers think about the mental states, deception, and social status), the extent to which the items made readers think about physical causality, and the rated imageability of the events of the story. None of these features explained variation in TPJ activation. The authors did not consider the number of action events within each stimulus as a potential confound, but a cursory analysis of the four examples given in the paper suggests a relatively tight positive correlation ($r = 0.98$) between TPJ activation level and number of events portrayed (Fig. 2). Though this result needs confirmation using the entire stimulus sample, it is not unexpected given other evidence relating processing of linguistic (verbs and event nouns) and nonlinguistic markers of events with activation in the posterior temporal and inferior parietal region (Bedny, Caramazza, Grossman, Pascual-Leone, & Saxe, 2008; Bedny, Dravida, & Saxe, 2014; Zacks, Speer, Swallow, & Maley, 2010). Thus, an alternative account of some of the evidence relating mentalizing to the TPJ is that the TPJ region processes event concepts, and that ToM stimuli used in some previous studies tended to contain a higher density of event concepts compared to control stimuli.

Summary

Our ability to learn about and interact with the world through acquisition, storage, retrieval, selection, analysis, and synthesis of concept representations is a defining feature of the human brain. The intent of this chapter was to point out how these core processes underlie various mental activities that use previously acquired knowledge. My central claim is that such activities, which include language use, remembering the past, planning and envisioning the future, reflecting on the self, making moral judgments, predicting and interpreting the behavior of others, and daydreaming, are all instances in which we retrieve and manipulate concepts. It would be ridiculous, of course, to conclude somehow from this account that the study of these specific kinds of conceptual processing is not valuable and worthwhile. Understanding the specific conceptual types and relationships that support a particular domain of knowledge processing is a central goal of cognitive science. Efforts to understand the neural correlates of these processing domains, including the domain of mentalizing, would benefit from a more explicit recognition of the general conceptual processes on which they rest, tighter experimental controls, and a more cautious attitude regarding claims of functional specificity.

References

Andrews-Hanna, J. R. (2012). The brain's default network and its adaptive role in internal mentation. *The Neuroscientist, 18*(3), 251–270.

Badre, D., Poldrack, R. A., Pare-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47*, 907–918.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11*(7), 280–289.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.

Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., & Saxe, R. (2008). Concepts are more than percepts: The case of action verbs. *Journal of Neuroscience, 28*, 11347–11353.

Bedny, M., Dravida, S., & Saxe, R. (2014). Shindigs, brunches, and rodeos: The neural basis of event words. *Cognitive, Affective, & Behavioral Neuroscience, 14*(3), 891–901.

Bedny, M., & Thompson-Schill, S. L. (2006). Neuroanatomically separable effects of imageability and grammatical class during single-word comprehension. *Brain and Language, 98*, 127–139.

Binder, J. R. (2016). In defense of abstract conceptual representations. *Psychonomic Bulletin and Review, 23*, 1096–1108.

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

Binder, J. R., Desai, R., Conant, L. L., & Graves, W. W. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*, 2767–2796.

Binder, J. R., McKiernan, K. A., Parsons, M., Westbury, C. F., Possing, E. T., Kaufman, J. N., & Buchanan, L. (2003). Neural correlates of lexical access during visual word recognition. *Journal of Cognitive Neuroscience, 15*(3), 372–393.

Binder, J. R., Medler, D. A., Desai, R., Conant, L. L., & Liebenthal, E. (2005). Some neurophysiological constraints on models of word naming. *NeuroImage, 27*, 677–693.

Binder, J. R., Westbury, C. F., Possing, E. T., McKiernan, K. A., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience, 17*(6), 905–917.

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*, 1–38.

Carreiras, M., Riba, J., Vergara, M., Heldmann, M., & Münte, T. F. (2009). Syllable congruency and word frequency effects on brain activation. *Human Brain Mapping, 30*, 3079–3088.

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *NeuroImage, 55*(2), 705–712.

Dove, G. O. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition, 110*, 412–431.

Fernandino, L., Binder, J. R., Desai, R. H., Pendl, S. L., Humphries, C. J., Gross, W., … Seidenberg, M. S. (2016). Concept representation reflects multimodal abstraction: A framework for embodied semantics. *Cerebral Cortex, 26*, 2018–2034.

Fliessbach, K., Wesi, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *NeuroImage, 32*, 1413–1421.

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.

Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., & Binder, J. R. (2010). Neural systems for reading aloud: A multiparametric approach. *Cerebral Cortex, 20*, 1799–1815.

Henson, R. N. A., Price, C. J., Rugg, M. D., Turner, R., & Friston, K. J. (2002). Detecting latency differences in event-related BOLD responses: Application to words versus nonwords and initial versus repeated face presentations. *NeuroImage, 15*(1), 83–97.

Husserl, E. (1973/1900). *Logical investigations* (J. N. Findlay, Trans.). London, England: Routledge.

Ingvar, D. H. (1985). Memory of the future: An essay on the temporal organization of conscious awareness. *Human Neurobiology, 4*, 127–136.

Ischebeck, A., Indefrey, P., Usui, N., Nose, I., Hellwig, F., & Taira, M. (2004). Reading in a regular orthography: An fMRI study investigating the role of visual familiarity. *Journal of Cognitive Neuroscience, 16*(5), 727–741.

Jessen, F., Heun, R., Erb, M., Granath, D. O., Klose, U., Papassotiropoulos, A., & Grodd, W. (2000). The concreteness effect: Evidence for dual-coding and context availability. *Brain and Language, 74*, 103–112.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex, 48*, 805–825.

Kotz, S. A., Cappa, S. F., von Cramon, D. Y., & Friederici, A. D. (2002). Modulation of the lexical-semantic network by auditory semantic priming: An event-related functional MRI study. *NeuroImage, 17*, 1761–1772.

Kuchinke, L., Jacobs, A. M., Grubich, C., Vo, M. L. H., Conrad, M., & Herrmann, M. (2005). Incidental effects of emotional valence in single word processing: An fMRI study. *NeuroImage, 28*, 1022–1032.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris, 102*(1–3), 59–70.

Mahy, C. E., Moses, L. J., & Pfeifer, J. H. (2014). How and where: Theory-of-mind in the brain. *Developmental Cognitive Neuroscience, 9*, 68–81.

Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology, 62*, 103–134.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature norms for a large set of living and nonliving things. *Behavior Research Methods, Instruments, & Computers, 37*(4), 547–559.

Mechelli, A., Gorno-Tempini, M. L., & Price, C. J. (2003). Neuroimaging studies of word and pseudoword reading: Consistencies, inconsistencies, and limitations. *Journal of Cognitive Neuroscience, 15*(2), 260–271.

Meteyard, L., Rodriguez Cuadrado, S., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex, 48*, 788–804.

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 211–277). New York, NY: McGraw-Hill.

Molenberghs, P., Johnson, H., Henry, J. D., & Mattingley, J. B. (2016). Understanding the minds of others: A neuroimaging meta-analysis. *Neuroscience and Biobehavioral Reviews, 65*, 276–291.

Nelson, D. L., & McEvoy, C. L. (2000). What is this thing called frequency? *Memory and Cognition, 28*, 509–522.

Orfanidou, E., Marslen-Wilson, W. D., & Davis, M. H. (2006). Neural response suppression predicts repetition priming of spoken words and pseudowords. *Journal of Cognitive Neuroscience, 18*(8), 1237–1252.

Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York, NY: Oxford University Press.

Prabhakaran, R., Blumstein, S. E., Myers, E. B., Hutchison, E., & Britton, B. (2006). An event-related fMRI investigation of phonological-lexical competition. *Neuropsychologia, 44*, 2209–2221.

Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.

Rissman, J., Eliassen, J. C., & Blumstein, S. E. (2003). An event-related fMRI investigation of implicit semantic priming. *Journal of Cognitive Neuroscience, 15*(8), 1160–1175.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Lawrence Erlbaum.

Sabsevitz, D. S., Medler, D. A., Seidenberg, M., & Binder, J. R. (2005). Modulation of the semantic system by word imageability. *NeuroImage, 27*, 188–200.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society of London: Series B, 362*(1481), 773–786.

Schilbach, L., Bzdok, D., Timmermans, B., Fox, P. T., Laird, A. R., Vogeley, K., & Eickhoff, S. B. (2012). Introspective minds: Using ALE meta-analyses to study commonalities in the neural correlates of emotional processing, social & unconstrained cognition. *PLoS One, 7*(2), e30920–e30920.

Schwanenflugel, P. (1991). Why are abstract concepts hard to understand? In P. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 223–250). Hillsdale, NJ: Erlbaum.

Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience, 21*(3), 489–510.

Sugiura, M., Sassa, Y., Watanabe, J., Akitsuki, Y., Maeda, Y., Matsue, Y., … Kawashima, R. (2006). Cortical mechanisms of person representation: Recognition of famous and personally familiar names. *NeuroImage, 31*, 853–860.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences USA, 94*, 14792–14797.

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping, 30*, 829–858.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., … Zilles, K. (2001). Mind reading: Neural mechanisms of theory of mind and self-perspective. *NeuroImage, 14*(1 Pt 1), 170–181.

Wallentin, M., Østergaarda, S., Lund, T. E., Østergaard, L., & Roepstorff, A. (2005). Concrete spatial language: See what I mean? *Brain and Language, 92*, 221–233.

Woodard, J. L., Seidenberg, M., Nielson, K. A., Miller, S. K., Franczak, M., Antuono, P., … Rao, S. M. (2007). Temporally graded activation of neocortical regions in response to memories of different ages. *Journal of Cognitive Neuroscience, 19*(7), 1113–1124.

Xiao, Z., Zhang, J. X., Wang, X., Wu, R., Hu, X., & Tan, L. H. (2005). Differential activity in left inferior frontal gyrus for pseudowords and real words: An event-related fMRI study on auditory lexical decision. *Human Brain Mapping, 25*, 212–221.

Zacks, J. M., Speer, N. K., Swallow, K. M., & Maley, C. J. (2010). The brain's cutting-room floor: Segmentation of narrative cinema. *Frontiers in Human Neuroscience, 4*, 168.