EMOTIONS AS SEMANTIC POINTERS: CONSTRUCTIVE NEURAL MECHANISMS

Paul Thagard and Tobias Schröder
University of Waterloo
Draft 6, March 2013

Thagard, P., & Schröder, T. (forthcoming). Emotions as semantic pointers: Constructive neural mechanisms. In L. F. Barrett & J. A. Russell (Eds.), *The psychological construction of emotions*. New York: Guilford.

INTRODUCTION

This chapter proposes a new neurocomputational theory of emotions that is broadly consistent with the psychological construction view (Russell, 2009; Gross and Barrett, 2011), and enhances it by laying out an empirically plausible set of underlying neural mechanisms. The new theory specifies a system of neural structures and processes that potentially explain a wide range of phenomena, supporting the claim that emotions are not just physiological perceptions or just cognitive appraisals or just social constructions. We claim that emotions can be understood as *semantic pointers*, a special kind of neural process hypothesized by Chris Eliasmith (2013) to provide explanations of many kinds of cognitive phenomena, from low-level perceptual abilities all the way up to high-level reasoning. Our aim is to show that Eliasmith's semantic pointer architecture for neural processing has the potential to account for a wide range of phenomena that have been used to support physiological, appraisal, and social constructionist accounts of emotion. We also discuss how it can be used to provide neural mechanisms for emotions as psychological constructions. Specifically, we will show how the semantic pointer hypothesis helps to specify how emotional states sometimes result from application of linguistic categories to representations of biological states, a core proposition of the psychological constructionist approach (e.g., Barrett, 2006; Barrett, Barsalou, & Wilson-Mendenhall, this volume; Lindquist & Gendron, 2013; Russell, 2009).

We begin with a quick review of the semantic pointer architecture (SPA), introducing its key ideas: neural representation, semantic pointers, binding, and control. These ideas have natural applications to emotions in ways that provide a unified account of their physiological, cognitive, and social aspects. Since SPA was designed as a framework for creating biologically plausible models of cognitive processes in general, its application to a theory of emotion is a contribution to "unifying the mind" (e.g., Barrett, 2009; Duncan & Barrett, 2007): The basic mechanisms governing emotion generation are no different from the mechanisms governing other aspects of human cognition. We sketch a neurocomputational model designed to show how SPA can be used to describe the neural mechanisms that generate emotional responses. We have not yet implemented this model computationally, but doing so should be straightforward using the NENGO simulation software that Eliasmith and his colleagues have developed (http://www.nengo.ca). Once implemented in a running computer program, the model provides a way of testing the semantic pointer theory of emotions through simulations of neural and psychological experiments. This methodology is common in cognitive science: A theory is a proposal about structures and processes, a model is a computational specification of these structures and processes, and a program is a running instantiation whose performance can be directly compared with results of experiments (Thagard, 2012a, ch. 1). We will also discuss the relevance of the new theory and model of emotions for issues about psychological construction and variability.

The Semantic Pointer Architecture

A cognitive architecture is a general proposal about the structures and processes that produce intelligent thought (Anderson, 1983, 2007; Newell, 1990; Thagard, 2012a).

It hypothesizes the kinds of mental representations and computational procedures that constitute a mechanism for explaining a broad range of kinds of thinking, including perception, attention, memory, problem solving, reasoning, learning, decision making, motor control, language, emotion, and consciousness. The most influential cognitive architectures that have been developed are either rule-based, using if-then rules and procedures that operate on them to explain thinking, or connectionist, using artificial neural networks (Thagard, 2005). Eliasmith (2013; Eliasmith et al., 2012) has proposed a different kind of cognitive architecture that is tied much more closely to the operations of the brain, but is capable of many of the kinds of high-level reasoning that make rule-based explanations plausible. We now present a highly simplified and informal account of the key ideas in his semantic pointer architecture; please consult his book for a much more mathematically rigorous and neurologically detailed description.

SPA adopts a view of neural representation and processing that is much more biologically accurate than the connectionist (parallel distributed processing, PDP) view of "brain-style" computing that has been highly influential in psychology since Rumelhart and McClelland (1986). SPA is based on the Neural Engineering Framework developed by Eliasmith and Anderson (2003). Like connectionist views, SPA assumes that mental representations are distributed processes involving the firing of populations of neurons connected by excitatory and inhibitory links. But SPA differs from connectionist models in at least the following ways:

1. Neurons spike, so that firing patterns and not just activations (firing rates) contribute to representational capacity.

2. Neurons are heterogeneous, with different temporal patterns resulting from different neurotransmitters.
3. Neural networks are large, involving thousands or millions of neurons rather than the few dozens typical of connectionist models.
4. Neural networks are organized into functional subnetworks, which may be mapped onto anatomical brain areas.

In sum, a mental representation according to SPA is a process involving the interaction of many thousands or millions of spiking neurons with varying dynamics.

Connectionism has been criticized as inadequate for explaining high-level cognitive phenomena such as reasoning and language use (e.g. Fodor and Pylyshyn, 1988; Jackendoff, 2002). SPA meets this challenge by showing how neural processes can support the kinds of problem solving and linguistic understanding usually attributed to symbolic rules. The key bridge from biologically realistic neurons to cognitive symbols is the idea of a semantic pointer. We think that this idea is the best available candidate for solving the symbol/subsymbol problem, one of the most important in cognitive science, which concerns the relations between high-level symbols such as words and concepts and low-level distributed representations in neural networks. In the context of a psychological constructionist perspective on emotion, the semantic pointer idea is important because it provides a detailed neural-level explanation of the conceptualization process thought to be central to the generation of specific emotions out of more blurry unspecific affective reactions (e.g., Barrett, 2006; Russel, 2009). More generally, we will argue that understanding emotions as semantic pointers allows one to integrate biological, cognitive, and socio-cultural constraints on emotional experience into one single framework.

Semantic pointers are neural processes that (1) provide *shallow meanings* through symbol-like relations to the world and other representations, (2) expand to provide *deeper meanings* with relations to perceptual, motor, and emotional information, (3) support complex syntactic operations, and (4) help to control the flow of information through a cognitive system to accomplish its goals. Thus semantic pointers have semantic, syntactic, and pragmatic functions, just like the symbols in a rule-based system, but with a highly distributed, probabilistic operation. Below we will show how emotions construed as semantic pointers have analogous semantic, syntactic, and pragmatic functions, while they are grounded in highly distributed representations of physiological states.

A semantic pointer consists of spiking patterns in a large population of neurons that provide a kind of compressed representation analogous to JPEG picture files or iTunes audio files. The term "pointer" comes from computer science where it refers to a kind of data structure that gets its value from a machine address to which it points. A semantic pointer is a neural process that compresses information in other neural processes to which it points and into which it can be expanded when needed. For example, the concept *car* can be understood as a semantic pointer consisting of spiking neurons that compress, point to, and expand into other populations of spiking neurons that contain a wide range of information in various modalities such as verbal and visual.

Eliasmith (2013) shows how symbolic pointers provide a mechanistic elucidation of the physical symbol system hypothesis of Barsalou (1999), according to which symbols are higher-level representations of combined perceptual components extracted from sensorimotor experience. For example, full understanding of the concept *car* requires both semantic knowledge about related concepts such as *machine* and *transportation*, and sensory and motor experiences such as *color*, *sound*, and *movement*, which are semantic primitives cannot be defined linguistically (Johnson-Laird, 1983; Johnson-Laird & Quinn, 1976). The semantic pointer *car* includes a compressed mental model of that motor experience.

The Semantic Pointer Architecture proposes a hierarchical organization of the cognitive system, where higher-level symbol-like representations point to neural representations at lower levels more tied to sensorimotor experience. The semantic pointer theory of emotion is thus compatible with the psychological constructionist approach, which also makes use of Barsalou's (1999) grounded cognition ideas (see Barrett, Barsalou, & Wilson-Mendenhall, this volume). To some degree, SPA is also compatible with views of embodied affect, cognition, and conceptual metaphor (e.g., Crawford, 2009; Lakoff & Johnson, 2003; Niedenthal, Winkielman, Mondillon, Vermeulen, 2009), while avoiding some of the more radical views according to which the mind does not employ representation or computation (Thagard, 2010b). SPA holds that meaning is often grounded in bodily experience; but semantic pointers are sufficiently decoupled from lower-level deep meanings (which they only represent in compressed form) that they can operate with symbol-like properties.

Why is this important for emotion? As we will describe in more detail below, such a hierarchical organization of the cognitive system allows for an integrative understanding of all the different facets of emotion, from the physiological components to socially constructed symbols. For example, the experience of *love* and *affection* is ultimately grounded in the sensory experience of physical warmth that infants of all cultures feel in the arms of their mothers (Lakoff & Johnson, 2003), but is also subject to considerable cultural symbolic construction and variation conveyed in the arts and literature over centuries (Belli, Harré, & Íñiguez, 2010). Accordingly, Osgood and colleagues have shown that the semantic relations between the culturally constructed symbols of language are constrained by the basic dimensions of emotion, evaluation-valence, potency-control, and activity-arousal (e.g., Fontaine, Scherer, Roesch, & Ellsworth, 2007; Heise, 2010; Osgood, May, & Miron, 1975; Osgood, Suci, & Tannenbaum, 1957; Rogers, Schröder, & von Scheve, forthcoming).

A major problem for connectionist cognitive architectures has been how to represent syntactically complex information such as relations (he loving her, which is different from she loving him) and rules (if she loves him, he is happy). The problem arose because early connectionist models were unable to represent such syntactic complexity, but various solutions have since been proposed. SPA adopts the solution proposed by Plate (2003) called holographic reduced representations that shows how vectors can be bound together in ways capable of maintaining syntactic complexity. A vector is a mathematical structure that represents 2 or more quantities; a vector with *n* quantities is called *n-dimensional*. If *man, loves,* and *woman* are all represented by vectors, then *loves(man, woman)* can be represented by another vector that binds the original ones by various mathematical operations including convolution, which interweaves structures. (For an introduction to holographic reduced representations and convolution, see Eliasmith and Thagard, 2001, and Thagard and Stewart, 2011.)

Representing symbolic concepts as vectors may sound technical, but it gains plausibility from a long tradition in the psychology of emotion. Specific emotions such as love or happiness have been represented as two-dimensional vectors in circumplex models of emotion (*pleasure* and *arousal*; Larsen & Diener, 1992; Russell, 1980) or as three-dimensional vectors in Osgood's affective space (*evaluation-valence*, *potency-control*, and *activity-arousal*; Osgood et al., 1957; 1975). More recently, Fontaine et al. (2007) mapped the Osgood space onto 144 single features of emotions (such as specific appraisals, physiological reactions, or muscle movements).

There is an analogy between these kinds of vectors and the SPA: symbol-like 3-dimensional representations of emotions (shallow meanings) can be viewed as compressed models of the underlying 144-dimensional representations of sensorimotor and appraisal correlates (deep meanings of emotion). Affect control theory (Heise, 2007), a mathematically formalized theory of social interaction and culturally shared emotional experience, represents social perception as vectors with nine or more dimensions, with the elements corresponding to Osgood's dimensions of affect (evaluation-valence, potency-control, and activity-arousal) as well as grammatical structure (agent, action, and recipient). The resulting mathematical model allows one to compute dynamic emotional states that occur during social interaction through combinations of vectors (Rogers, Schröder, & von Scheve, forthcoming).

Eliasmith has shown that these kinds of vector structures and operations can be accomplished by populations of spiking neurons. Hence SPA networks are capable of all the syntactic complexity needed for many kinds of reasoning and language processing. Eliasmith et al. (2012) have used SPA to create a large scale model of the functioning brain, whose 2.5 million neurons can perform tasks like symbol recognition and reproduction, memory, and question answering. SPA can also be used to solve the famous Tower of Hanoi problem that has long been a benchmark of rule-based inference (Stewart and Eliasmith, 2011), to explain behavioral priming (Schröder & Thagard, 2013), and to model the interplay between intentional and impulsive action (Schröder, Stewart, & Thagard, forthcoming). In the present chapter, we argue that SPA as a general cognitive architecture is also a candidate for understanding the neural mechanisms underlying the psychological construction of emotion, in line with the view that "affect is [also] a form of cognition" (Duncan & Barrett, 2007).

Ignoring the mathematics, we can say that a neural architecture can represent relations such as *loves(man, woman)* by producing a nested binding like this one:

loves(man, woman) = BIND[BIND(loves, action), BIND(man, agent), BIND(woman, recipient)].

That is, the relation *loves* is represented by a binding of three separate bindings of elements: the binding of loves as the action, the binding of the man as the agent, and the binding of the woman as the recipient. Each of the elements can be represented by a vector and hence translated (using Eliasmith's method) into a pattern of activity in a population of neurons. The bindings themselves can be understood both as mathematical operations on the vectors and as neural processes that transform neural processes into more complex ones. Semantic pointers are important for this process because they provide the high-level representation of concepts such as *love*.

In addition to semantics and syntax, a cognitive architecture must support pragmatics, the ability of a system to accomplish its goals by making decisions in particular contexts. Eliasmith (2013) describes the need for a system to manage the control of: selection of which representations to employ, manipulation of current representations given current context, determination of what next course of action to pursue, and so on. SPA has various mechanisms for managing these kinds of control, corresponding to operations in the basal ganglia and other brain regions. Let us now see how SPA ideas about semantics, syntax, and pragmatics apply to emotions.

EMOTIONS AS SEMANTIC POINTERS

SPA is sufficiently broad that it could be used to provide neural mechanisms for any of the currently popular approaches to understanding emotions. The view that emotions are primarily perceptions of physiological states, espoused by Damasio (1994), Prinz (2004) and others, could be specified biologically by means of neural populations that respond to physiological variables such as heart rate, skin response, and hormone levels. This view is compatible with the claim that there are innate mechanisms for basic emotions such as happiness, fear, and anger (Ekman, 2003). Alternatively, the view that emotions are cognitive appraisals of the relevance of a situation to an agent's goals, espoused by Oatley (1992), Scherer, Schorr, and Johnstone (2001), and many others, could be specified biologically by means of neural populations that compute the goal relevance of a situation. The radical view of Harré (1989) and others that emotions are mere social constructions created by an individual's culture could be specified biologically by means of neural populations that respond to social communication. As mentioned before, the psychological constructionist view that emotional states emerge from a conceptualization process, where linguistic categories are applied to make sense of inner representations of physiological states, can be specified through the compression and binding mechanisms of semantic pointers which connect deep, sensorimotor meanings with higher-level symbolic meanings.

Much more interestingly, SPA can be used to build a model that synthesizes physiological, appraisal, social, and psychological constructionist accounts of emotions, in line with the EMOCON account of emotional consciousness (Thagard and Aubie, 2008). On this account, physiological and appraisal theories of emotions are not alternatives, but can be unified by a neural model that shows how parallel processing in the brain can integrate both perception of physiological states and evaluation of the relevance of the current situation to the goals of the agent. Figure 1 shows some of the relevant brain areas, with the amygdala and insula most relevant for physiological perception, and frontal areas most relevant to cognitive appraisal.

EMOCON was too large and complex to be implemented computationally by the tools available when it was conceived in 2006, but SPA is supported by software that should make possible an enhanced models using new theoretical ideas such as semantic pointers. EMOCON is consistent with the view that there are social and cultural aspects of emotion, because appraisal is subject to social coordination through the interpersonal (often nonverbal) communication of affect (Manstead & Fischer, 2001; Parkinson & Simons, 2009). Furthermore, goals and beliefs relevant for the appraisal process are often culturally acquired. Let us examine how the general structure of EMOCON can be worked out in much more biological detail via the theoretical resources of SPA. Later we will sketch a new model, POEM, that includes the brain areas in figure 1 plus additional ones identified as relevant to emotions by brain scanning experiments.

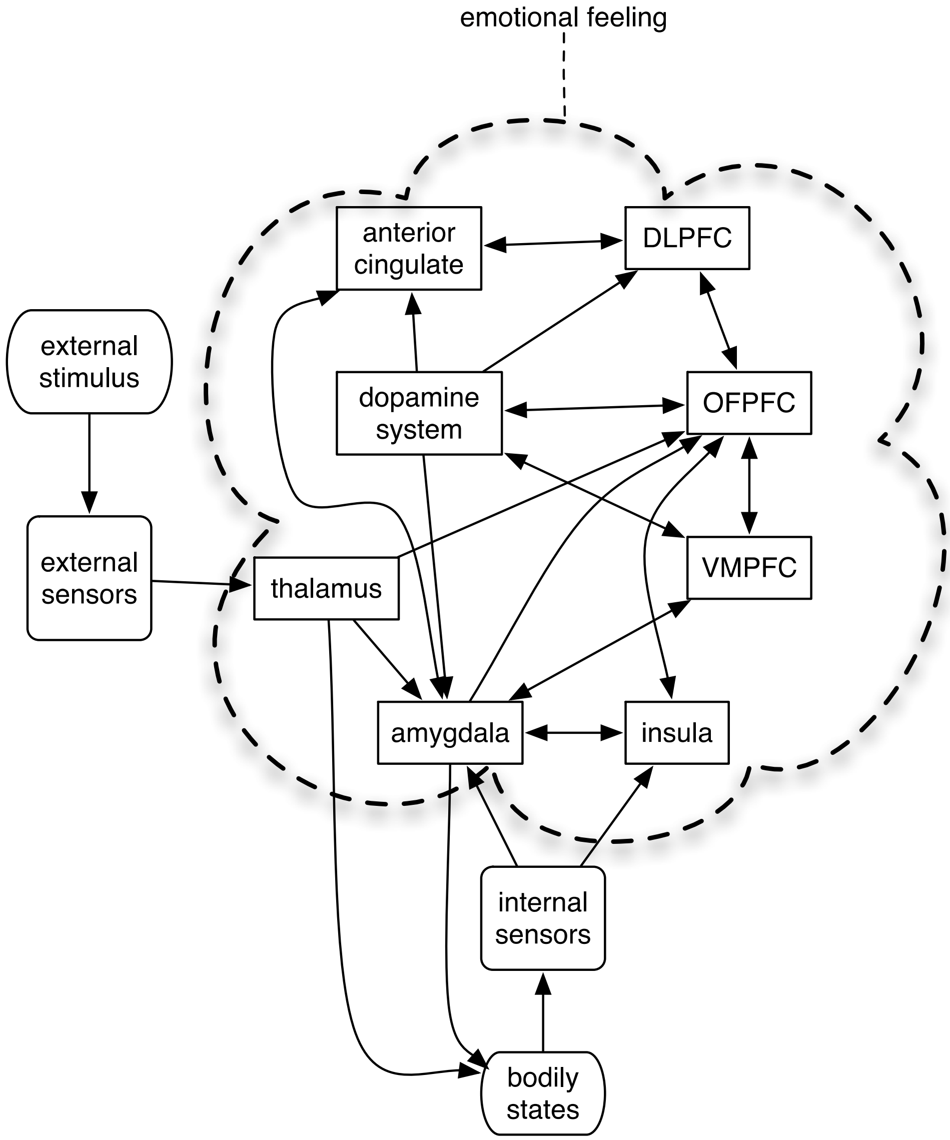

Figure 1. The EMOCON model of Thagard and Aubie (2008), which contains details and partial computational modeling. Abbreviations: DLPFC is dorsolateral prefrontal cortex. OFPFC is orbitofrontal prefrontal cortex. VMPFC is ventromedial prefrontal cortex. The dotted line is intended to indicate that emotional consciousness emerges from activity in the whole system.

We need to spell out how the four main ideas of SPA – neural representation, semantic pointers, binding, and control – apply to emotions. To begin, we need to distinguish between emotion tokens, such as particular instances of happiness or fear, and emotion types, which are classes of emotional responses. For now, we are only concerned with emotion tokens, and will address the question of types below. We propose to identify instances of emotions with SPA-style neural representations consisting of spiking behavior in neural populations. This proposal assumes the contentious philosophical mind-brain identity theory, defended elsewhere (Thagard, 2010a).

It is crucial to note that a neural population need not be confined to a specific brain region, because there are extensive synaptic connections between neurons in different regions. We need to avoid the naïve assumption that particular emotions reside in particular brain regions, for example fear in the amygdala and happiness in the nucleus accumbens, because brain scans find that emotions are correlated with much more distributed brain activity (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012). The EMOCON model in figure 1 assumes that emotions are widely distributed across multiple brain areas, and a SPA account of emotions can similarly look at emotions as activity in populations of neurons not confined to a single region.

Now we get to the crucial contention of this chapter, that an emotion token is a semantic pointer. We need to show how emotions can possess shallow meanings through compressed representations that can be expanded into deeper representations, employing binding operations that support syntactic combinations and manage the control of the flow of information in a system capable of generating actions that accomplish goals.

What is going on in your brain when you feel an emotion such as being happy upon hearing that a paper has been accepted by a good journal? This occurrence of emotion consists of firing of interconnected neurons in a way that amounts to a compressed but expandable representation of the current situation. The compressed representation functions to produce verbal reports of the experience – "I'm happy" – as well as to link the incident to other occurrences of feeling happy. Moreover, this neural representation "points" to others that provide an expanded meaning for it is based on *both* physiological perception and cognitive appraisal, in accord with the EMOCON model.

Binding operations of the sort that can be performed by convolution in SPA are crucial for the construction of the emotional response of being happy that your paper was accepted. In logical notation, we could represent the situation by this formula:

accept (journal, paper). P1

Linguistically, the relevant proposition is actually more complicated, since it should also contain the information that the paper is yours, but we ignore the self until the next section. The neural processing of P1 requires a binding something like:

BIND[BIND(accept, action), BIND(journal, agent), BIND(paper, recipient)]. P2

In SPA, this complex binding can be understood as mathematical operations on vectors that are performed by neural populations, producing a new vector that is also represented by a firing pattern in neurons. P2 is the proposition (vector, neural firing pattern) that results from accomplishing the bindings that unite the representations of acceptance, journal, and paper into a proposition representing an event.

Now we have to find a neural representation for the proposition that you are happy that the paper was accepted. Verbally, this is just something like:

happy(accept(journal, paper)). P3

But if happiness results from a combination of physiological perception and cognitive appraisal, then the binding is more complicated, something like:

BIND (P2, physiology, appraisal). P4

In P4, P2 is the firing pattern already produced by binding together *happy, journal,* and *paper*, and the other two ingredients need to be explained.

We can easily identify the physiological and appraisal aspects of the emotional response as firing patterns in neural populations. For physiology, the relevant process is detection by brain areas including the amygdala and insula of such bodily changes as heart rate, breathing rate, skin temperature, configuration of facial muscles, and hormone levels. For appraisal, the relevant process is a complex calculation of goal relevance that can perform constraint satisfaction, as in the EACO (Emotional Appraisal as Coherence) computational model in Thagard and Aubie (2008). The journal acceptance presumably satisfies a variety of personal goals such as increasing your reputation and salary. Binding these three ingredients together – the situation representation P2, the representation of physiological changes, and the representation of appraisals – produces the semantic pointer that constitutes the neural representation of your emotional reaction. We are not using "bind" as a vague metaphor here, because SPA contains the required algorithms for using neural populations to perform such bindings. Some emotions may also involve a self-representation, discussed below.

To summarize, figure 2 provides a rough picture of how an emotion token can be a semantic pointer through operations of compression and convolution. The groups of circles depict populations of thousands or millions of neurons. The semantic pointer is the result of binding of several key representations: the situation that the emotional response is about, the results of physiological perceptions, and the results of cognitive appraisal. Each of these three representations could be further decompressed into other representations. Usually, the situation representation is a pointer to a binding of an action with an agent. The physiological representation would be a compression of many kinds of changes in bodily states such as breathing. The cognitive appraisal representation would be a compression of a complex parallel satisfaction process that determines the goal-significance of the situation. Hence the occurrence of emotions depends on a cascade of semantic pointers pointing to others, and emotion representations have rich meanings deriving from their relations to other representations and ultimately to perceptual inputs. Emotion representations in the form of semantic pointers are also able to contribute to complex syntactic structures via binding, making possible highly structured propositions such as "If you're happy and you know it, clap your hands."

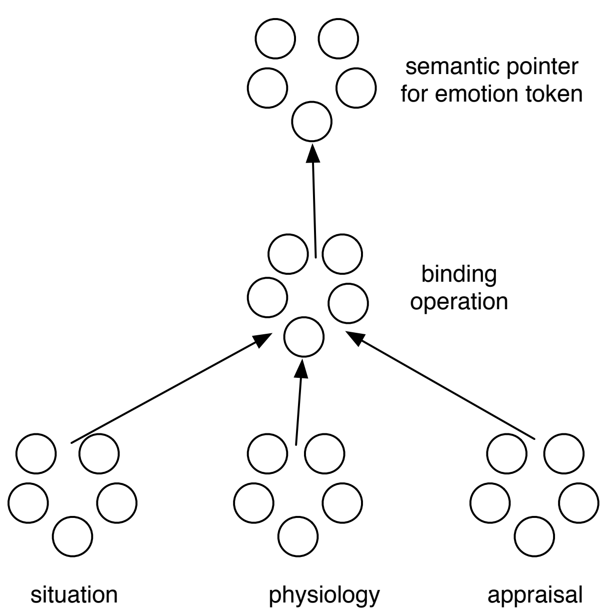

Figure 2. Schematic view of how a semantic pointer for an emotion occurrence compresses the binding of a situation, physiological reaction, and cognitive appraisal. The circles represent large and diverse neural populations, while the arrows represent transformations of neural activity in those populations through interactions with other populations.

Semantic pointers as in figure 2 can represent complex emotions such as being happy about a journal acceptance, and can help to control the flow of information in your brain, form intentions, and hence lead to action. The representation that you are happy will affect your behavior in various ways, including how you will react to other people. Internally, the positive feelings associated with the journal, the paper, yourself, and the people you collaborated with will increase, leading to thoughts that may enhance your subsequent interactions with them. Elsewhere, we use the semantic pointer hypothesis to explain the influences of emotional states on automatic and intentional actions (Schröder, Stewart, and Thagard, forthcoming; Schröder & Thagard, 2013). Moreover, the case can be made that all human cognition is controlled in part by emotion (Clore and Palmer, 2009), so the semantic pointer theory of emotions is potentially an important part of *any* cognitive theory (see also Duncan & Barrett, 2007). Because semantic pointers connect to control and hence to action in particular contexts, they have important contributions to pragmatics as well as semantics and syntax.

Understanding emotions as semantic pointers has several theoretical advantages. First, it applies the Semantic Pointer Architecture, which provides a powerful synthesis of symbolic, connectionist, and embodiment approaches. Second, it shows how constructionist views of emotions can meld with biological accounts. Third, it provides a precise way of integrating physiological and appraisal views of emotion within a neural framework that is sufficiently rigorous to be amenable to computer simulation. In order to incorporate social construction accounts, we need to say more about self-representation.

SELF-REPRESENTATION AND SOCIAL CONSTRUCTION

The last section used P2 to represent your happiness that your paper was accepted by a journal, but left out the "you". P2 needs to be expanded into something like:

happy(you, accept(journal, paper)). P5

This says that you are happy that the journal accepted the paper, but how can we represent you? There is a voluminous literature in social psychology on the nature of the self, much of it concerned with self-representations such as self-concepts (e.g. Thagard in press). Here we offer the novel view that a self-representation is a semantic pointer, i.e. a neural representation that expands contextually into much richer multimodal representations of sensory experiences, bodily states, emotional memories, and social conceptions.

Once again, binding is the key idea for tying together diverse representations. Hume (1888) worried that the self was no more than a bundle of perceptions, whereas Kant (1965) and other dualistic philosophers have sought the unity of the self in transcendental entities like souls. Binding provides the clue to a semi-unified, naturalistic view of self-representation as involving the integration of several factors by convolution and other neural transformations, along the following lines:

Self-representation = BIND(self-concepts, experiences, memories). P6

Here self-concepts are themselves semantic pointers for general concepts that people apply to themselves, including roles such as *university professor*, *father*, and *colleague* and traits such as *tall*, *middle-aged*, *sociable*, and *conscientious*. Through making use of such concepts, people incorporate cultural knowledge and social structure into the current representation, since the meaning of roles and traits is culturally constructed and passed on across generations through language (Berger & Luckmann, 1966). MacKinnon and Heise (2010) have argued that language provides humans with an implicit cultural theory of people. Whenever one applies linguistic categories to make sense of oneself, one internalizes part of the institutional structure of society and its set of behavioral and emotion-related expectations that are crystallized in the language (Heise, 2007; MacKinnon & Heise, 2010).

The social constructionist interpretation of being happy upon acceptance of one's journal article is that the emotion is a consequence of the culturally constructed meanings of the concepts included in that representation (cf. Rogers, Schröder, & von Scheve, forthcoming). With our semantic pointer theory of emotion, we intend to provide a detailed explanation of how such culturally constructed semantics are constrained by, and related to the biological processes in the brain (for a more general discussion of social vs. psychological constructionist approaches, see Boiger & Mesquita, this volume). All linguistic concepts include affective meanings (Osgood, May, & Miron, 1975), which are widely shared among members of one culture (Heise, 2010; Moore, Romney, Hsia, & Rusch, 1999). From our perspective, affective meanings are compressed representations of the corresponding appraisal patterns and physiological reactions that are bound into the complex representation of the emotion. Hence the decompressing mechanisms of semantic pointers enable culturally shared symbolic concepts to activate and socially synchronize related cognitive and physiological processes in the generation of emotion and action (for details, see Schröder & Thagard, 2013).

In accord with the Semantic Pointer Architecture operating within the Neural Engineering Framework, both the process of binding and the structures bound are patterns of spiking activity in populations of neurons. Then, the complex proposition P5 that you are happy that the journal accepted your paper becomes the kind of nested, compressed structure shown in figure 3.

Not all emotional experiences require a self-representation. The journal acceptance case does, where it is important that it was *your* paper that was accepted for publication and the emotional response is immediately identifiable as yours. But there are many species that seem to have emotional reactions similar to those of humans but have no detectable sense of self. Members of only eight species are currently known to be able to recognize themselves in mirrors: humans, gorillas, chimpanzees, orangutans, elephants, dolphins, pigs, and magpies (Prior, Schwarz, and Güntürkün, 2008; Broom, Sena and Moynihan, 2009). Yet behaviorally and neurologically, emotions such as fear seem to operate in very similar ways in a variety of mammals such as rats (LeDoux, 1996). We conclude, therefore, that self-representations and their linguistic expressions are not essential to emotions, even though they are an important part of human emotions.

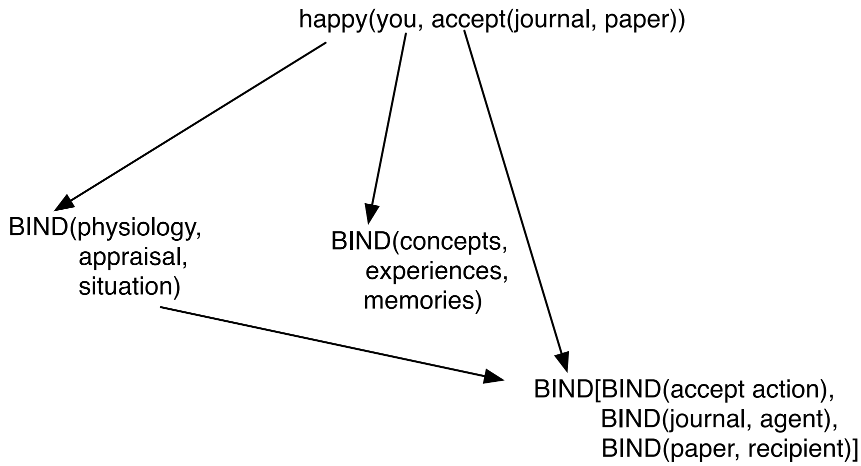

Figure 3. Representation of the proposition that you are happy that your paper was accepted as the semantic pointer to representations that bind emotion, self, and situation. The representation of you (the self) is the binding of concepts, experiences, and memories.

Self-representations are particularly important to social emotions such as pride, guilt, shame, and envy, which require an appreciation of one's location in a social situation. Pride, for example, usually involves a positive feeling of accomplishment as appreciated by a social group one cares about, such as a family or a profession. Such emotions clearly have a large cultural component, because different societies attach wildly different values to behaviors such as work, education, and gender roles. Depending on the broad values accepted in a society, substantially different behaviors can generate different social emotions. In this sense, emotions are socially constructed, but only partially, because underlying them are the same biological mechanisms, common to all people: representation, binding, and control operating in brain areas that collectively accomplish physiological perception and cognitive appraisal. The neuroanatomy is common to all humans, but the results of appraisal will vary dramatically depending on what are viewed by individuals in a particular culture to be the appropriate goals.

Wierzbicka (1999) surveys the commonalities of emotions across many cultures as well as the extensive diversity. Our semantic pointer account of emotions can account for such findings: while the neural mechanisms akin to all humans are the source of cross-cultural commonality, more complex semantic pointers such as in our example of article acceptance are sufficiently decoupled from underlying bodily representations that they are influenced by culturally-bound patterns of social interpretation.

Cultural influences on emotion are clearly not confined to high-level social emotions such as pride. Anger, fear, disgust, happiness, sadness, and surprise occur broadly (Ekman, 2003), but the evaluation of particular objects as scary, disgusting, and so on varies with a culture's beliefs and attitudes about the objects. Hence all human emotions have important social dimensions, but they are not mere social constructions because of their substantial underlying biological commonalities.

PSYCHOLOGICAL CONSTRUCTION

We are now in a position to try to answer the key questions posed by the editors of this volume. Our boxes are short to avoid duplicating material already presented.

1. **What are the key ingredients from which emotions are constructed? Are they specific to emotion or are they general ingredients of the mind? Which, if any, are specific to humans?**

Emotions are constructed from neural processes that involve semantic pointers, binding, and control, operating on multiple brain areas to integrate physiological perception and cognitive appraisal. The basic neural processes are common to all psychological operations, but the integration operations are specific to emotion. We speculate that the social emotions are restricted to humans because they require a high-level, linguistic representation, but that other emotions are common in other mammals.

2. What brings these ingredients together in the construction of an emotion? Which combinations are emotions and which are not (and how do we know)?

The various physiological and cognitive ingredients in the construction of an emotion are brought together by general neural mechanisms of binding and compression. Because of the substantial amount of evidence that both physiological perception and cognitive appraisal are relevant to emotions, it is reasonable to conjecture that emotions differ from other psychological processes in that they are combinations of neural representations of *both* physiology and appraisal. In addition, for a few species such as humans, self-representations can be bound into the overall emotion.

3. How important is variability (across instances within an emotion category, and in the categories that exist across cultures)? Is this variance epiphenomenal or a thing to be explained?

Because of the large degree of variation in the beliefs, attitudes, and goals across different cultures, we should expect some degree of variability in emotion categories, as suggested by the linguistic evidence. This variance can be explained by the role that societies play in inculcating the cognitive-affective elements, especially attitudes and goals that are crucial for cognitive appraisal. Despite this variance, there is also the need to explain apparent commonalities deriving from biological universals in brain anatomy. Little evidence is available to speak to the question of how much variability there is within particular emotion categories within particular societies.

BUILDING AND TESTING A MODEL

The ideas we sketched about how the semantic pointer architecture can generate an integrative theory of emotions that still needs to be fleshed out by a much more fully developed computational model. We now present the design of such a model, called "POEM" for "POinters EMotions", to be implemented using the software platform Nengo which incorporates the theoretical ideas of Eliasmith's Neural Engineering Framework and Semantic Pointer Architecture.

First, POEM will specify neural populations corresponding to the relevant brain areas, which will include all the ones in the EMOCON model (figure 1) as well as additional ones such as the anterior temporal lobe identified as important for emotions by various researchers (Lindquist et al., 2012). We will need to decide what dynamics are appropriate for the different brain areas. For example, areas that are part of the dopamine pathway such as the nucleus accumbens and the orbitofrontal prefrontal cortex may have different temporal characteristics than other areas that employ different neurotransmitters. We will need to determine whether these temporal differences matter for modeling psychological phenomena.

Second, we will use neuroanatomical information to determine the interconnectivity of the brain areas selected, establishing links between the neurons in them.

Third, we will ensure that the resulting network can perform the following functions: input from external perception of information about a situation; input and interpretation of perceptual information from internal, bodily sources; appraisal of the situation information with respect to goals; and integration of all these kinds of information in an overall interpretation of the significance of the situation, generating a broadly distributed network whose firing patterns will be identifiable as corresponding to particular emotional responses.

Fourth, we will test the POEM model by seeing whether it is capable of simulating the results of important psychological experiments. The previous EMOCON model was never tested in this way, because available technology did not support a computational implementation and because the main explanatory target of the model was emotional consciousness, about which little experimentation has taken place. In contrast, earlier neurocomputational models of emotion were strenuously tested for the ability to duplicate experimental results. The GAGE model of Wagar and Thagard (2004) was used to simulate the behavior of participants in the Iowa gambling task experiments of Bechara, Damasio, Tranel, and Damasio (1997) and the behavior of participants in the famous attribution experiments of Schacter and Singer (1962). The ANDREA model of Litt, Eliasmith, and Thagard (2008) was used to simulate the behavior of experiments based on decision affect theory (Mellers, Schwartz, & Ritov, 1999) and prospect theory (Kahneman, & Tversky, 1979). Both of these models were used to duplicate not only the qualitative behavior of human participants in the relevant experiments, but also the quantitative behavior, generating results close to the results found by the experimenters. The theoretical assumptions behind POEM are highly compatible with those behind

GAGE and ANDREA, so it should be feasible to incorporate the structure of those models into a broader one that includes new ideas about binding, semantic-pointers, and self-representation.

Ideally, POEM will be able to match GAGE and ANDREA in their ability to simulate important experimental results about emotions, and will also be able to simulate quantitatively the results of additional key experiments from different paradigms such as embodiment (e.g., Niedenthal et al., 2009), appraisal theory (e.g., Siemer & Reisenzein, 2007), and sociology of emotion (e.g., Heise & Weir, 1999). In addition, POEM should be able to incorporate the ability of EMOCON to provide a mechanistic explanation of how conscious experiences of emotions arise by incorporating *competition* among semantic pointers (Thagard, forthcoming). Building POEM is a daunting task, but within the range of the convenient software tools of Nengo and the powerful theoretical ideas of NEF and SPA.

As usual in cognitive science, the proof is in the programming. Success in building POEM will demonstrate the feasibility of semantic pointers, binding, and control for providing a mechanistic account of emotions. Progress in simulating the results of a broad variety of psychological and neurological experiments will provide evidence that the set of mechanisms incorporated into POEM correspond approximately to those that produce emotions in human brains. In addition to applying POEM to psychological experiments, we will need to validate its neurological assumptions by comparing its structure and performance to the neuroscientific evidence including brains scans using fMRI and other techniques.

4. What constitutes strong evidence to support a psychological construction to emotion? Point to or summarize empirical evidence that supports your model or outline what a key experiment would look like. What would falsify your model?

Strong evidence to support a Semantic Pointer Architecture interpretation of the psychological construction of emotions would come from its ability to simulate the results of a wide variety of experiments concerning emotions such as Schachter and Singer (1962). While building POEM, we will compile a list of key experiments that will provide targets for simulation. We will strive to show that POEM behaves similarly to experiment participants at the quantitative as well as qualitative levels. Contrary to Karl Popper's philosophy of science, direct falsification of theory by evidence is not an important part of scientific practice (Thagard 1988, 2010a). Rather, theories compete to explain the evidence, and a theory is rejected when it is surpassed by another with greater explanatory power. The POEM model of emotions would be falsified if other researchers produce a model that explains more experimental results.

DISCUSSION

Earlier we made the claim that emotion tokens (particular occurrences of emotions such as a person feeling happy at a particular time) are semantic pointers, a special kind of compressed neural pointer that decompresses into a binding of information about physiological states, cognitive appraisals, and sometimes the self. It would probably be just as reasonable to claim that an emotion token is the *combined* process that includes both the semantic pointer and the neural processes to which it points. Given the imprecise nature of folk psychological categories such as *happy,* more specific identification is not supportable by evidence. What matters is that we now have a plausible candidate for a neural mechanism that can have the various psychological effects produced by specific occurrences of emotions.

Even more problematic is the identification of the whole category or type of happiness and other emotions with a class of neural processes. There is insufficient evidence about the neural correlates of kinds of emotions to specify an identity relation between a whole class of occurrences of kinds of emotions and a whole class of neural processes. Hence we support a token-token mind-brain identity theory of emotions, while remaining agnostic concerning the viability of a type-type identity theory for emotions. There are indeterminacies on both sides: not only is there insufficient neurological evidence about kinds of emotions such as happiness, there is still uncertainty about whether categories derived from folk psychology will have a legitimate role in a well worked out theory of emotions supported by experimental evidence from both psychology and neuroscience. As Patricia Churchland (1986) suggested, we should not take folk categories for granted, but should expect a co-evolution of psychological and neural theories that may substantially modify ideas such as happiness.

There are recent technological advances that may prove to be useful for establishing type-type identities. One is a novel brain mapping approach that combines text mining and meta-analysis to enable accurate and generalizable classification of cognitive states (Yarkoni, Poldrack, Nichols, Van Essen, and Wager, 2011). In lexical brain decoding, the text of a large corpus of articles is retrieved and a search string for a psychological state such as "pain" serves to retrieve a subset of articles that report neural coordinates. An automatic meta-analysis of the coordinates produces a whole-brain map of the probability of the psychological state given activation at each brain location.

Another potentially promising technique is the application of machine learning algorithms that can detect specific distributed activation patterns in the brain which reliably correspond to the mental representation of specific concepts. This method has proven useful for identifying the neural correlates of intentions (Haynes, Sakaai, Rees, Gilbert, Frith, & Passingham, 2007), cognitive tasks (Poldrack, Halchenko, & Hanson 2010), and tool concepts (Shinkareva, Mason, Malave, Wang, Mitchell, & Just, 2008). The combination of brain imaging and machine learning has been used for diagnosing eating disorders (Weygandt, Schaefer, Schienle, & Haynes, forthcoming). In principle, the challenge is no different for associating emotions with neural activity. Recognition of the locations that correlate generally with psychological states such as pain or happiness, along with a theoretical account of the neural processes that connect activities in those locations, may lead to reasonable type-type identifications of kinds of emotions with kinds of neural processes. We leave open the possibility, however, that such studies will provide grounds for revising everyday concepts of emotions.

We avoid pursuing philosophical issues about emotions that are not addressable by empirical methods. For example, both the question of whether emotions supervene on physical states and the question of whether emotions have essences (one common interpretation of the claim that emotions are natural kinds) presuppose a conception of necessity, which most philosophers understand as truth in all possible worlds. Minds supervene on brains if and only if *necessarily* a difference in mental properties requires a difference in neural properties. An essence of something is a property that it has necessarily. It is hard enough to collect evidence concerning what emotions are in *this* world, let alone to pursue the impossible task of collecting evidence that would address questions about what they are in all possible worlds. To modify Wittgenstein, whereof one cannot collect evidence, thereof one must be silent. See Thagard (2010a) for a more thorough defense of philosophical naturalism in discussions of mental states, including rejection of the metaphysical concept of necessity. Hence the concepts of supervenience, essence, and natural kind are best ignored in science-oriented discussions of emotions. Whether the general category of emotions (and the myriad categories of different types of emotions) survive in scientific discourse will depend on the roles they play in neuropsychological theorizing.

We have tried to contribute to such theorizing by outlining a new, neurocomputational theory of the neural construction of emotions. This theory is broadly compatible with psychological construction views that do not tie emotions to localized, programmed neural operations. Instead, we propose that emotions as semantic pointers are compressed representations of bindings of physiological perceptions and cognitive appraisals, operating in many different brain areas. The plausibility of this theory will depend on the success of the planned model POEM in explaining, more effectively than alternative models, the full range of psychological and neural evidence about emotions.

Acknowledgments. Paul Thagard's work is supported by the Natural Sciences and Engineering Research Council of Canada. Tobias Schröder was awarded a research fellowship by the Deutsche Forschungsgemeinschaft (# SCHR 1282/1-1) to support this work.

REFERENCES

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (2007). *How can the mind occur in the physical universe?* Oxford: Oxford University Press.

Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review, 10*, 20-46.

Barrett, Lisa. 2009. The future of psychology: Connecting mind to brain." *Perspectives on Psychological Science* 4: 326-39.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-660.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*, 1293-1295.

Belli, S., Harré, R., & Íñiguez, L. (2010). What is love? Discourse about emotions in social sciences. *Human Affairs: A Postdisciplinary Journal for Humanities and Social Sciences, 20*, 240-270.

Berger, P. L. & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*, Garden City, NY: Anchor Books.

Broom, D. M., Sena, H., & Moynihan, K. L. (2009). Pigs learn what a mirror image represents and use it to obtain information. *Animal Behavior, 78*, 1037-1041.

Churchland, P. S. (1986). *Neurophilosophy*. Cambridge, MA: MIT Press.

Clore, G. L. (2009). Affective guidance of intelligent agents: How emotion controls cognition. *Cognitive Systems Research, 10*, 22-30.

Crawford, L. E. (2009). Conceptual metaphors of affect. *Emotion Review, 1*, 129-139.

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: G. P. Putnam's Sons.

Duncan, S., & Barrett, L. F. (2007). Affect is a form of cognition: A neurobiological analysis. *Cognition & Emotion, 21*, 1184-1211.

Ekman, P. (2003). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*. New York: Henry Holt.

Eliasmith, C. (2013). *How to build a brain*. Oxford: Oxford University Press.

Eliasmith, C., & Anderson, C. H. (2003). *Neural engineering: Computation, representation and dynamics in neurobiological systems*. Cambridge, MA: MIT Press.

Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., and Rasmussen, D. (2012). A large-scale model of the functioning brain. *Science, 338*, 1202-1205.

Eliasmith, C., & Thagard, P. (2001). Integrating structure and meaning: A distributed model of analogical mapping. *Cognitive Science, 25*, 245-286.

Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*(3-81).

Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science, 18*, 1050–1057.

Gross, J. J., & Barrett, L. F. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review, 3*, 8-16.
Harré, R. (Ed.). (1989). *The social construction of emotions*. Oxford: Blackwell.
Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R. E. (2007). Reading hidden intentions in the brain. *Current Biology, 17*, 323-328.
Heise, D. R. (2007). *Expressive order: Confirming sentiments in social action.* New York: Springer.
Heise, D. R. (2010). *Surveying cultures. Discovering shared conceptions and sentiments*. New York: Wiley.
Heise, D. R., & Weir, B. (1999). A test of symbolic interactionist predictions about emotions in imagined situations. *Symbolic Interaction, 22*, 139-161.
Hume, D. (1888). *A treatise of human nature*. Oxford: Clarendon Press.
Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
Johnson-Laird, P. N. & Quinn, J. G. (1976). To define true meaning. *Nature, 264*, 635-636.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263-291.
Kant, I. (1965). *Critique of pure reason* (N. K. Smith, Trans. 2 ed.). London: MacMillan.
Lakoff, G. & Johnson, M. (2003). *Metaphors we live by (with a new afterword)*. (2. Ed.). Chicago: University of Chicago Press.
Larsen, R. J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. In M. S. Clark (Ed.), *Review of personality and social psychology* (Vol. 13, pp. 25-59). Thousand Oaks: Sage.
LeDoux, J. (1996). *The emotional brain*. New York: Simon and Schuster.
Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences, 35*, 121-143.
Lindquist, K. A., & Gendron, M. (2013). What's in a word? Language constructs emotion perception. *Emotion Review, 5*, 66-71.
Litt, A., Eliasmith, C., & Thagard, P. (2008). Neural affective decision theory: Choices, brains, and emotions. *Cognitive Systems Research, 9*, 252-273.
MacKinnon, N. J. & Heise, D. R. (2010). *Self, identity, and social institutions*. New York: Palgrave Macmillan.
Manstead, A. S. R., & Fischer, A. H. (2001). Social appraisal: The social world as object of and influence on appraisal processes. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal Processes in Emotion: Theory, Research, Application* (pp. 221-232). New York: Oxford University Press.
Mellers, B., Schwartz, A., & Ritov, I. (1999). Emotion-based choice. *Journal of Experimental Psychology: General, 128*, 332-345.
Moore, C. C., Romney, A. K., Hsia, T.-L., & Rusch, C. D. (1999). The universality of the semantic structure of emotion terms: Methods for the study of inter- and intra-cultural variability. *American Anthropologist, 101*, 529-546.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Niedenthal, P. M., Winkielman, P., Mondillon, L., Vermeulen, N. (2009). Embodiment of emotion concepts. *Journal of Personality and Social Psychology*, 96(6), 1120-1136.

Oatley, K. (1992). *Best laid schemes: The psychology of emotions*. Cambridge: Cambridge University Press.

Osgood, C. E., May, W. H., & Miron, M. S. (1975). *Cross-cultural universals of affective meaning.* Urbana: University of Illinois Press.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The Measurement of Meaning.* Urbana: University of Illinois Press.

Parkinson, B. & Simons, G. (2009). Affecting Others: Social Appraisal and Emotion Contagion in Everyday Decision Making. *Personality and Social Psychology Bulletin*, 35, 1071-1084.

Plate, T. (2003). *Holographic reduced representations*. Stanford: CSLI.

Poldrack, R. A., Halchenko, Y., & Hanson, S. J. (2010). Decoding the large-scale structure of brain function by classifying mental states across individuals. *Psychological Science, 20*, 1364-1372.

Prinz, J. (2004). *Gut reactions: A perceptual theory of emotion*. Oxford: Oxford University Press.

Rogers, K. B., Schröder, T., & von Scheve, C. (forthcoming). Dissecting the sociality of emotion: A multi-level approach. *Emotion Review*.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161-117

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition & Emotion, 23*, 1259-1283.

Schacter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69*, 379-399.

Scherer, K. R., Schorr, A., & Johnstone, T. (2001). Appraisal processes in emotion: Theory, Methods, Research. New York: Oxford University Press.

Schröder, T., Stewart, T. C., & Thagard, P. (forthcoming). Intention, emotion, and action: A neural theory based on semantic pointers. *Cognitive Science*.

Schröder, T., & Thagard, P. (2013). The affective meanings of automatic social behaviors: Three mechanisms that explain priming. *Psychological Review, 120*, 255-280.

Shinkareva, S. V., Mason, R. A., Malave, V. L., Wang, W., Mitchell, T. M., & Just, M. A. (2008). Using fMRI brain activation to identify cognitive states associated with percpetion of tools and dwellings. *PLOS One*, e1394.

Siemer, M., & Reisenzein, R. (2007). The process of emotion inference. *Emotion, 7*, 1-20.

Stewart, T. C., & Eliasmith, C. (2011). Neural Cognitive Modelling: A Biologically Constrained Spiking Neuron Model of the Tower of Hanoi Task. *33rd Annual Conference of the Cognitive Science Society, Austin, TX: Cognitive Science Society (2011)*.

Thagard, P. (1988). *Computational philosophy of science*. Cambridge, MA: MIT Press.

Thagard, P. (2005). *Mind: Introduction to cognitive science* (2nd ed.). Cambridge, Mass.: MIT Press.

Thagard, P. (2010a). *The brain and the meaning of life*. Princeton, NJ: Princeton University Press.

Thagard, P. (2010b). How brains make mental models. In L. Magnani, W. Carnielli & C. Pizzi (Eds.), *Model-based reasoning in science and technology. Abduction, logic, and computational discovery* (pp. 447-461). Berlin: Springer.

Thagard, P. (2012a). Cognitive architectures. In K. Frankish & W. Ramsay (Eds.), *The Cambridge handbook of cognitive science*. Cambridge: Cambridge University Press.

Thagard, P. (2012b). *The cognitive science of science: Explanation, discovery, and conceptual change*. Cambridge, MA: MIT Press.

Thagard, P. (forthcoming). Creative intuition: How EUREKA results from three neural mechanisms. In L. M. Osbeck & B. S. Held (Eds.), *Rational intuition: Philosophical roots, scientific investigations* Cambridge: Cambridge University Press.

Thagard, P. (in press). The self as a system of multilevel interacting mechanisms. *Philosophical Psychology*.

Thagard, P., & Aubie, B. (2008). Emotional consciousness: A neural model of how cognitive appraisal and somatic perception interact to produce qualitative experience. *Consciousness and Cognition, 17*, 811-834.

Thagard, P., & Stewart, T. C. (2011). The Aha! experience: Creativity through emergent binding in neural networks. *Cognitive Science, 35*, 1-33.

Wagar, B. M., & Thagard, P. (2004). Spiking Phineas Gage: A neurocomputational theory of cognitive-affective integration in decision making. *Psychological Review, 111*, 67-79.

Weygandt, M., Schaefer, A., Schienle, A., & Haynes, J.-D. (2012). Diagnosing different binge-eating disorders based on reward-related brain activation patterns. *Human Brain Mapping, 33*, 2135-2146.

Wierzbicka, A. (1999). *Emotions across languages and cultures: Diversity and universals*. Cambridge: Cambridge University Press.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods, 8*, 665-670.