Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 58 (2019) 35–53

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

The semantic pointer theory of emotion: Integrating physiology, appraisal, and construction ☆

Action editor: Eva Hudlicka

Ivana Kajić [a], Tobias Schröder [b,*], Terrence C. Stewart [a], Paul Thagard [a]

[a] University of Waterloo, Canada
[b] Potsdam University of Applied Sciences, Germany

Received 18 July 2018; received in revised form 4 March 2019; accepted 24 April 2019
Available online 2 May 2019

Abstract

Emotion theory needs to explain the relationship of language and emotions, and the embodiment of emotions, by specifying the computational mechanisms underlying emotion generation in the brain. We used Chris Eliasmith's Semantic Pointer Architecture to develop POEM, a computational model that explains numerous important phenomena concerning emotions, including how some stimuli generate immediate emotional reactions, how some emotional reactions depend on cognitive evaluations, how bodily states influence the generation of emotions, how some emotions depend on interactions between physiological inputs and cognitive appraisals, and how some emotional reactions concern syntactically complex representations. We contrast our theory with current alternatives, and discuss some possible applications to individual and social emotions.

Keywords: Emotion; Language; Appraisal; Construction; Multi-level mechanisms; Affective computing; Neural engineering framework; Semantic pointers

Theorizing about emotions began with Aristotle, but there is still no generally accepted account of what emotions are and how they work. From a biological perspective, emotions are physiological states that enable an organism to quickly adjust behavior to the needs of the body (Craig, 2002; Damasio & Carvalho, 2013; James, 1884). From a cognitive perspective, emotions are rapid mechanisms for appraising the significance of events in the environment with respect to an individual person's goals (Oatley & Johnson-Laird, 1987; Ortony, Clore, & Collins, 1990; Scherer, Schorr, & Johnstone, 2001). From a sociological perspective, emotions help to control social interactions and sustain the social and cultural order (Heise, 2007; Hochschild, 1983; Kemper, 2006; Von Scheve, 2014). Most emotion research has relied on experimental approaches and verbal theorizing, which have generated comparatively little insight into the dynamic mechanisms of emotion (Reisenzein et al., 2013; Scherer, 2009). Studying the computational dynamics of emotion is not only of theoretical value, but also has practical implications for affective computing, a subfield of artificial intelligence that seeks to build virtual agents that interact

☆ Author note: Authorship is alphabetical due to equal contributions. Tobias Schröder did part of this work as a postdoctoral fellow at the University of Waterloo, supported by the German DFG (SCHR1282/1-1). Ivana Kajić was a graduate student at the Bernstein Center for Computational Neuroscience in Berlin during early stages of the work. She acknowledges current support by the AFOSR (FA9550-17-1-0026). Paul Thagard received funding from the Natural Sciences and Engineering Research Council of Canada (NSERC).
* Corresponding author at: Potsdam University of Applied Sciences, Kiepenheuerallee 5, 14469 Potsdam, Germany.
E-mail address: post@tobiasschroeder.de (T. Schröder).

https://doi.org/10.1016/j.cogsys.2019.04.007
1389-0417/

smoothly with humans (Gratch, Cheng, & Marsella, 2015; Gratch & Marsella, 2004; Hoey, Schröder, & Alhothali, 2016; Picard, 1997; Reisenzein et al., 2013; Scherer, BAdnziger, & Roesch, 2010).

A cognitive architecture is a general proposal about the representations and computations that produce intelligent thought. Current cognitive architectures – rule-based, connectionist, and Bayesian – have had substantial success in explaining important aspects of thinking such as problem solving, learning, and categorization. A cognitive architecture is a general theory of the mechanisms of thought, proposing parts (representations) and interactions (computations) that aim to explain the full range of mental phenomena. It is much broader than models of particular aspects of cognition such as memory, learning, and emotion.

Applications of cognitive architectures to emotion have been rare despite their evident potential for theoretical coherence (cf. Reisenzein et al., 2013). Existing computational models of emotion have largely been concerned with cognitive appraisal, although a few have attended to embodied aspects of emotions (Balkenius et al., 2016; Hudlicka, 2011; Marsella & Gratch, 2014, 2016). Integrating the physiological, cultural, and linguistic levels of explanation, but also more general cognitive processes such as decision making, social interaction, memory, language, attention, perception, and creativity remains a challenge. This wider cross-domain knowledge integration is important because there is now substantial psychological evidence that emotions are integral to all of these processes (Barrett, Lewis, & Haviland-Jones, 2016). Moreover, neuroscientific evidence has mounted that emotion mechanisms in the brain are thoroughly integrated with cognitive mechanisms (Damasio, 1994; Duncan & Barrett, 2007; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012; Pessoa, 2013).

We aim to show that Chris Eliasmith's (2013) Semantic Pointer Architecture (SPA) provides the basis for a broad and rigorous theory of emotions. We propose that neural mechanisms involved in processing of emotions can be described as *semantic pointers*: patterns of neural firing that bind neural representations of physiological inputs, evaluations of situations, and cultural/linguistic context. Spaun, a computer model of the brain based on SPA, consists of millions of simulated neurons that can perform tasks such as symbol recognition, categorization, memory storage and retrieval, and motor control (Eliasmith et al., 2012). Integrating emotions into SPA is an important step towards a general computational model of the mind (cf. Barrett, 2009; Lindquist & Barrett, 2012; Reisenzein et al., 2013).

After describing the operation of semantic pointers and their relevance to emotions, we present POEM (for POinters-EMotions), a neurocomputational model that simulates major empirical phenomena related to emotional experience. We conclude by discussing the relations of our model to major contemporary approaches to emotion research, including basic emotions (Ekman & Cordaro, 2011), appraisal theory (Ortony et al., 1990; Scherer et al., 2001), psychological constructionism (Barrett, 2017; Barrett & Russell, 2014; Lindquist et al., 2012; Russell, 2009), and the sociology of emotions (Hochschild, 1983; Lively & Heise, 2004, 2014).

What are emotions? Rather than attempting the hopeless task of defining "emotion" using necessary and sufficient conditions, we can characterize emotions by providing standard examples, typical features, and explanatory uses (Blouw, Solodkin, Thagard, & Eliasmith, 2016; Fehr & Russell, 1984; Thagard, 2019a). The exemplars of emotions include happiness, sadness, fear, anger, disgust, surprise, shame, embarrassment, pride, and so on. Typical features of these emotions include physiological changes, cognitive judgments, social influences including linguistic ones in humans, neural patterns in multiple brain areas, and enjoyable or painful experiences. Emotions help to explain people's actions, verbal reports, and conscious experiences.

In cognitive science, explanatory theories usually consist of descriptions of mechanisms that can generate important phenomena (Thagard, 2012b). Mechanisms are combinations of interconnected parts whose interactions produce regular changes. In theoretical neuroscience, the parts are neurons with synaptic connections that produce firing patterns through excitatory and inhibitory interactions.

Our theory of emotion consists of the following hypotheses. 1. Emotions are semantic pointers, which are patterns of firing in spiking neurons that integrate information of different sorts by neural bindings. 2. The information bound into semantic pointers for emotions includes external stimuli, physiological changes, stored concepts, and linguistic knowledge. 3. Emotional reactions occur when stimuli generate semantic pointers that combine physiological perception, cognitive appraisal, and social context. 4. Emotions become conscious when semantic pointers cross a threshold by outcompeting other semantic pointers.

We demonstrate the explanatory power of the semantic pointer theory of emotions by using the computational model POEM to simulate six important aspects of emotions: 1. Some stimuli generate immediate emotional reactions. 2. Some emotional reactions depend on cognitive evaluations of external stimuli. 3. Bodily states influence the generation of emotions. 4. Some emotions depend on interactions between physiological inputs and cognitive appraisal. 5. Some emotional reactions concern syntactically complex representations. 6. Sometimes, people can have ambivalent reactions to events, resulting in mixed emotions.

1. The Semantic Pointer Architecture (SPA)

To portray semantic pointers, we first specify them as the result of neural mechanisms involving binding of neural firing rates. We then characterize them mathematically and indicate their computational functions. Finally, we contrast semantic pointers with more familiar accounts of mental representation. We provide an elementary exposition of the Semantic Pointer Architecture because its components are relevant to explaining important aspects of emotions. Representations using groups of spiking neurons explain how brains can have the many varieties and objects of emotions. Binding by convolution explains how the brain combines neural representations of situations, cognitive appraisals, and physiological perceptions into semantic pointers for emotions.

1.1. Biological mechanisms

In everyday use, a representation is something that stands for something else, for example when the word "cat" or a drawing of a cat stands for cats in the world. In cognitive theory, a mental representation is a structure in the mind that similarly stands for things, for example when a word-like concept stands for a class of things or when a sentence-like proposition stands for a state of affairs in the world such as that cats have legs.

The development of cognitive science provided new ways of thinking about mental processes using ideas drawn from computation and neuroscience. Computationally, a representation is a structure in a programming language such as a number or string of letters that can be manipulated by algorithms to produce new structures, for example when a computer calculates that $2 + 4 = 6$. Biologically, a representation is the activity of a group of neurons that can be used to produce new representations through the interactions of the neurons. Let us look at this special kind of computation in more detail.

A neuron is a cell capable of receiving chemical inputs in the form of neurotransmitters and hormones, building up an electrical charge, and then firing (spiking) when the charge passes a threshold. Firing sends outputs to other neurons when the electrical signal releases neurotransmitters that can then affect the firing of connected neurons.

A single neuron can be a very simple representation when it becomes tuned to fire in response to sensory inputs from the world, as when a neuron fires in response to retinal signals. For example, a neuron could represent red if it fired rapidly when a red stimulus is presented, but slowly otherwise. But neurons have the capability of not only firing fast or slow, but also of firing in patterns, for example in the difference between FIRE REST FIRE REST and FIRE FIRE REST REST, which have the same firing rate but different patterns of firing, and there are far more firing patterns than firing rates. Hence a single neuron has a huge representational capacity, for mathematical reasons described in the following section.

Representations in the brain, however, rarely use single neurons, because groups of neurons working together have even greater representational capacity. Hence groups of neurons working together can stand for more complicated aspects of the world such as a cat with three white legs.

To go from the representation of simple features to more complicated situations, the brain must accomplish the binding of features, for example the binding of face and body and legs into cat, and the binding of cat and black into black cat. How do firing patterns in groups of neurons accomplish such binding? One popular theory is that binding results from different neurons becoming synchronized so that they fire with similar patterns (Singer, 2007), but this theory has difficulty explaining how there can be bindings of bindings of bindings, needed to produce more complicated representations such as the black cat snarling at the white cat (for discussion of binding see Feldman, 2013).

In the Semantic Pointer Architecture, binding is accomplished by a mathematical operation called circular convolution. Circular convolution is performed by neurons that take two patterns of firing and produce a new pattern of firing that can involve some of the same neurons. In turn, this new pattern can then be bound with other patterns to produce more and more complicated representations if there are enough neurons to carry out repeated bindings. An important property of convolution as a method of binding is that it can be undone, at least approximately, so that the combination of white and leg into white leg can be taken apart to yield an approximation to the original neural patterns for white and leg.

Semantic pointers are representations that are produced by binding of two or more patterns of neural firing that are semantic in two ways. First, they retain connections to sensory inputs from the world. The pattern of firing that represents cat results from convolutions that include a pattern for leg that originates perceptually through visual or tactile perceptions of legs, so the semantic pointer for cat retains some connection with the world. The term "pointer" indicates that such representations point to their sensory origins because convolution can be undone to approximate to the perceptually-derived inputs that went into it.

The second way in which semantic pointers are semantic is that they show how neural representations can be related to each other in the same way that linguistic symbols are. The mental representation *cat* is not just perceptually related to the world, but also to other representations as in the beliefs that cats eat mice and that some cats are pets. Such higher representations also can be produced by convolution, for example binding *cat*, *eat*, and *mice* into *cat eat mice*. Thus semantic pointers bridge the gap between neural network cognitive architectures and ones based on rules using symbols.

The Semantic Pointer Architecture has employed semantic pointers in simulations with millions of artificial neurons to model numerous important cognitive phenomena such as visual pattern recognition, concept application, motor control, and social priming (Blouw et al., 2016; Crawford, Gingerich, & Eliasmith, 2016; Eliasmith, 2013; Eliasmith et al., 2012; Rasmussen, Voelker, & Eliasmith, 2017; Schröder & Thagard, 2013). Before developing a semantic pointer theory of emotion, we will flesh out the idea of semantic pointers mathematically and comparatively.

1.2. Mathematics and computation

Let us consider a neuron with a firing rate of around 100 times per second, which gives it 100 possible firing rates but 2^{100} possible firing patterns over a second, which is more than 1 followed by 30 zeros. Ten neurons working together with each firing 100 times per second can have $10^{(2^{100})}$ firing patterns over a second, which is more than the number of atoms in the universe. Hence the representational capacity of groups of neurons firing in coordination with each other is enormous.

Learning changes the connectivity of groups of neurons by adjusting the strengths of the synaptic links between them in response to inputs from the environment, disposing the neurons to fire in appropriate patterns when similar inputs are received. But patterns of firing should be able to respond not only to simple features but also to combinations of features, so that concepts like *cat* and *pet* can substantially increase an organism's ability to understand the world.

Mathematically, patterns of firing in neural groups can be used to represent vectors, which are strings of numbers that represent a collection of dimensions. Vectors are commonly used to represent empirically measurable properties of objects and events in an environment. For example, the speed and direction of a car can be represented by the two-dimensional vector (20, 90), where 20 indicates that the car is going 20 km per hour and 90 indicates an angle of direction. Speed and location could be represented by a 3-dimensional vector that includes map coordinates, for example if (20, 10, 20) indicates the same speed and map indicators of 10 units across and 20 up.

Spiking neurons can also represent such vectors. For every vector dimension, the contribution of every neuron to the value in that dimension is computed by weighting the firing rate of that neuron. Such weights can be computed analytically for a known range of values in that dimension (e.g., maximal and minimal speed, or maximal and minimal angle). In SPA, maximal firing rates are randomly selected to simulate the neuronal variety observed in biological brains. To achieve a precise representation of a numerical value with neurons, it is common for the number of neurons in a group to exceed the number of dimensions of a vector the group is representing, so that multiple neurons contribute to a representation of a number.

Once a vector is represented by the firing pattern of a group of neurons, we can define convolution as a mathematical operation on vectors. The addition of vectors is simple, for example when adding (0.2, 0.4, 0.6, 0.8) and (0.1, 0.1, 0.1 0.1) yields (0.3, 0.5, 0.7, 0.9). But vector addition and multiplication of vectors could not perform convolution, because they obliterate the information in the input which cannot be reconstructed even approximately. The vector (0.3, 0.5, 0.6, 0.9) could have resulted from the addition of many other pairs of vectors.

Convolution solves this problem by using a more complex function that wraps the dimensional values of two vectors around each other to produce a vector that is not similar to the original vectors but can nevertheless be decomposed into an approximation of the originals. For details, see Eliasmith (2013, p. 406). In SPA, the convolution of vectors is computed by feed-forward neural networks, thereby implementing binding by convolution in spiking neurons.

Semantic pointers formed by convolution are a key component of a general account of neural computation called the Semantic Pointer Architecture that produces motor outputs in response to visual inputs and numerous other processes including working memory and action selection. This architecture is implemented in a computational model called Spaun that has been used to simulate many cognitive processes and even to control robots (Eliasmith et al., 2012).

1.3. Comparisons

Semantic pointers can be further characterized by contrasting them with more familiar theoretical ideas in cognitive science, including distributed representations, rules, schemas, modality-specific simulations, Bayesian inference, and somatic markers.

1.3.1. Distributed representations

In connectionist architectures such as the highly influential PDP account, concepts are not represented by single word-like nodes but rather by groups of neurons working together: the representation is distributed across multiple neurons as occurs in the brain (Rogers & McClelland, 2004; Rumelhart, McClelland, & Group, 1986). Semantic pointers are a kind of distributed representation, but they differ from standard accounts in two ways. First, the neurons employed in SPA are spiking neurons, capable of firing with many different patterns that affect the firing activity of other neurons. In contrast, PDP models use rate neurons, whose activity is determined by how fast they fire, not their particular patterns. Second, whereas PDP connections are formed by training using algorithms such as backpropagation, semantic pointers are formed by convolution of other neural patterns, including ones formed by learning. This formation enables semantic pointers to be subject to further convolutions into sentence-like representations, functioning as symbols.

1.3.2. Rules

In rule-based architectures such as ACT (adaptive control of thought), rules are IF-THEN structures where the IF and THEN parts are word-like symbols (Anderson, 1983, 2007). Semantic pointers are not rules, but they can be used to model rule-like problem solving, for example in the Tower of Hanoi problem (Eliasmith, 2013, chap. 5). The IF and THEN parts in SPA are not traditional symbols, but are represented by vectors created in part by convolution in ways that can be encoded as patterns of neural firing.

1.3.3. Schemas

In psychology, a schema is a mental representation of a class of objects events or practices. Semantic pointers provide a mechanistic explanation of what schemas are and how they work in the brain. For example, the restaurant schema is a combination of a concept and expectations for what to do in restaurants. Concepts can be explained as semantic pointers (Blouw et al., 2016), and expectations can be modeled as rules of the sort just described. Hence schemas in the brain are one kind of semantic pointer, but there are other kinds of semantic pointers such as representations of individuals.

1.3.4. Modality-specific simulations

Barsalou (2016) has emphasized that concepts are grounded in particular modalities such as vision rather than being amodal, language-like representations. Such concepts can be used in simulations, for example when visual images associated with a restaurant allow people to imagine themselves being seated and fed. Semantic pointers provide neural explanations of how concepts can combine different modalities and can serve in simulations using non-verbal rules. Hence Barsalou's psychological theories are mechanistically explained by SPA.

1.3.5. Bayesian inference

Bayesian cognitive architectures explain human thought using network structures that make inferences using probability theory. SPA does not explicitly use probabilities, but semantic pointer vectors can be interpreted as probability distributions, and Bayesian inference is approximately implemented by neural computations (Eliasmith, 2013, p. 281).

1.3.6. Somatic markers

According to Damasio (1994), cognitive processes such as decision making are heavily influenced by neural responses to bodily changes. These somatic markers enable the brain rapidly to evaluate and anticipate different actions and outcomes. Somatic markers are naturally understood as semantic pointers if they involve bindings of inputs from various internal physiological sensors and signals concerned with situations. Hence somatic markers are part (but not all) of the semantic pointer theory of emotion.

In sum, semantic pointers are processes in groups of spiking neurons that provide representations by repeatedly binding sensory, motor, and verbal information. Such processes provide a neural explanation of many psychological ideas such as schemas, and they are naturally extended to emotions.

2. Emotions in the Semantic Pointer Architecture

Emotions integrate multiple levels of meaning from embodied experiences through culturally constructed linguistic structures. For example, the experience of love is grounded in the sensory experience of physical warmth that infants of all cultures feel in the arms of their parents (Lakoff et al., 2003). However, the meaning of love is also shaped by the cultural history of a society as conveyed in art and literature over centuries, creating substantial cross-cultural variation (Belli, Harré, & Iñiguez-Rueda, 2010). The degree of universality versus cultural specificity may vary across different categories of emotions. Some emotions, like sadness and anger, have been considered more basic and universal (Oatley, 2009; Oatley & Johnson-Laird, 1987), while others, like shame or jealousy, are more conceptually and socially complex and thus more contingent on culture-specific language.

Semantic pointers are well equipped to combine the embodiment and cultural aspects of emotions. They are compatible with views of embodied affect, cognition, and conceptual metaphor (Crawford, 2009; Lakoff et al., 2003; Niedenthal, Winkielman, Mondillon, & Vermeulen, 2009); but they avoid the more radical view that the mind does not employ representation or computation (cf. Thagard, 2012a; Thagard & Schröder, 2014). We propose that emotions are semantic pointers that bind information that includes external stimuli, physiological changes, stored concepts, and linguistic knowledge.

Consider for example a situation where you are excitedly falling in love with a person named Pat. Your neural representation of Pat may combine verbal, visual, tactile, auditory, and even olfactory information. Your physiological response to Pat could include many changes such as increased heart rate, rapid breathing, sexual excitement, and elevated levels of neurotransmitters and hormones such as dopamine, opioids, cortisol, and oxytocin. You may also consciously or unconsciously evaluate Pat as meeting your social and other needs, where the evaluation combines both universal human needs such as belonging and relatedness with cultural expectations about what contacts are allowed and desired. Semantic pointers show how your brain can combine all of these into a unitary neural representation that is felt as love. Schematically, loving Pat = (bind (bind verbal, visual, etc.) (bind heart, breathing, hormones, etc.) (bind cultural expectations, goal appraisal, etc.)). This summary representation is sufficiently powerful to frequently become conscious because it outcompetes other representations such as daily events. Moreover, it influences actions by interacting with intentions to produce motor behaviours (cf. Schröder, Stewart, & Thagard, 2014).

The mathematical interpretation of semantic pointers as vectors naturally carries over to emotions, thanks to affect control theory, a mathematically formalized social psychological theory of social interaction and emotion (Heise, 2007; Hoey et al., 2016; Lively & Heise, 2014; Rogers, Schröder, & von Scheve, 2014; Schröder, Hoey, & Rogers, 2016). This theory treats social concepts denoting types of persons, traits, actions, and social settings as vectors in a three-dimensional EPA space defined by evaluation (goodness versus badness), potency (powerfulness versus powerlessness), activity (liveliness versus torpidity) (Osgood, May, & Miron, 1975). These vectors are obtained from empirical studies in which respondents rate hundreds of concepts with the semantic differential technique developed by Osgood. Speakers of a common language generally agree on the placement of social concepts in EPA space, so that these ratings can be interpreted as reflecting a within-culture consensus on the affective meaning of human sociality (Ambrasat, von Scheve, Conrad, Schauenburg, & Schröder, 2014; Heise, 2010). Emotion concepts such as happiness also have a representation in the EPA space because of their ratings for evaluation, potency, and activity. Fontaine, Scherer, Roesch, and Ellsworth (2007) have linked not just emotion concepts, but also the specific physiological features and appraisal patterns associated with those concepts to the EPA space.

Affect control theory models emotion generation in specific situations by taking as input the EPA vectors of concepts used to describe a situation and returning as output a new vector representing the emotional response to the situation (Heise, 2007; Hoey et al., 2016). The concept vectors are derived from empirical studies that elicit EPA ratings of linguistic event descriptions (e.g. "a professor who yells at a student") and compare them to EPA ratings of the constituents of these descriptions (e.g., "a professor", "to yell at someone", "a student") (see Heise, 2010, for details). A computer model called INTERACT (Heise, 1997) generates situation-specific emotion predictions, corresponding to cultural emotion norms, such as that a professor who yells at a student is likely angry. We show below how the computations for emotion generation by affect control theory can be performed by spiking neurons in the Semantic Pointer Architecture. The result is a naturalistic explanation of how social norms embedded in language constrain the generation of emotions in the brain.

3. POEM: a neurocomputational model of emotion

We evaluate the explanatory power of the semantic pointer theory of emotions by showing that a computational model derived from it can simulate numerous aspects of emotions. POEM (for POinter-EMotions) is implemented using the open-source neural network simulator Nengo (Stewart, Tripp, & Eliasmith, 2009). Nengo is used to create neural networks for simulation of biological and cognitive phenomena using the programming language Python, giving users the ability to control the level of technical detail pertaining to the network. Fig. 1 shows an example of a small neural network built in Nengo together with the code that generated that network. The network consists of two groups of neurons, one represents a number, and the other one represents a square of that number. The desired function (in this case it is a square of a number) is specified as a parameter of the connection that connects the two groups of neurons. In the background, Nengo uses the Neural Engineering Framework (NEF) to describe how groups of neurons can represent numerical values and to compute weights that yield desired computations over those values (Eliasmith et al., 2003).

POEM consists of such multiple interconnected groups of neurons labelled as: external sensation, internal perception, language, episodic memory, concepts, EPA space, and executive control. The activity of neurons (more specifically, their *firing patterns*) in each group is used to represent real-valued vectors and in total all groups contain about 300,000 neurons. The activity of some groups can be controlled externally by dictating which vector should be represented by that group. Such control is used when we want to present inputs to the model – by assigning labels to random vectors, we create a vocabulary where each concept has its unique vector representation. For example, the concept "cat" might be encoded as (0.1, 0.4, −0.5), and the concept "dog" as (−0.7, 0.9, 0.5). Then, if we want a group of neurons to represent "cat" we can set the activity of that

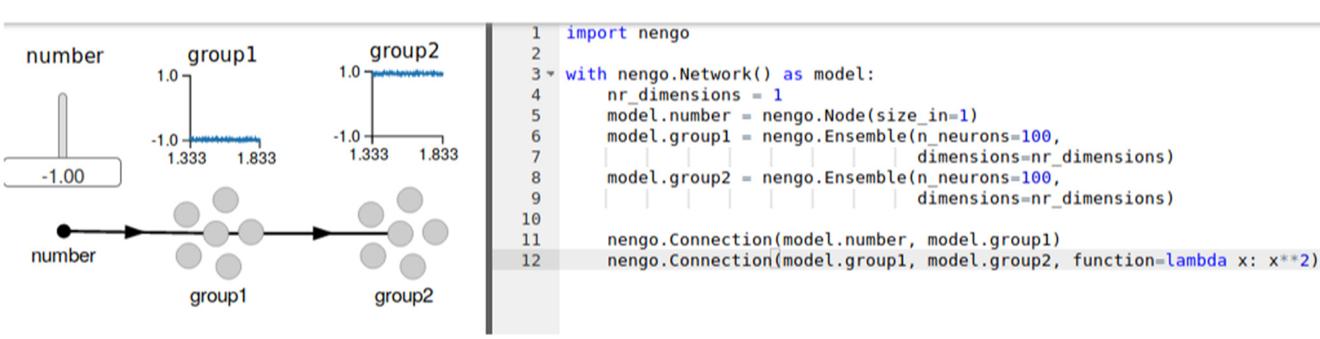

Fig. 1. Graphical User Interface of the Nengo neural network simulator. The Python code used to build the neural network that computes square of a number is shown on the right hand-side in the text editor, while the visual depiction of the generated network is on the left hand-side.

group to correspond to the vector for “cat” using the graphical user interface in Nengo.

In POEM we use 512-dimensional random vectors to represent 55 different concepts used in simulations of six different emotional phenomena. Some concepts, such as “snake” or “cake” are used to set the activity of the group representing external sensation. More specifically, if we set a group of neurons to represent the vector for “snake” we would interpret this as simulating a situation where a snake is encountered in an environment in some form (e.g., the sound of a snake hissing, the image of a snake in the distance). Concepts such as “joy” or “fear” are used to describe the emotional response, and such responses are observed as vectors represented by the executive control group of neurons. In the same way in which we can set the activity of a group of neurons to represent a vector, we can also interpret the activity of a group as representing a vector. That is, we can read out the vector represented by that group and compare how similar that vector is to all the vectors in our vocabulary (vectors in the vocabulary can be seen representing an “ideal” pattern). We use cosine similarity to compute the similarity between two vectors. The resulting similarity score between the represented vector and the ideal pattern is used to express how strongly the group of neurons represents a concept. If the score is high, we can interpret that as the group of neurons representing that vector. If the score is low, we can infer that the vector is not being represented by that group, and if it is anything in between we can interpret that as the level of similarity.

POEM implements a Leaky-Integrate-and-Fire (LIF) spiking neuron model, which is the default neuron model in Nengo. LIF models the change of neuronal activity over time, thus allowing us to simulate sequence-dependent events. As described in the previous sections, the activity of spiking neurons forms firing patterns that are used to represent vectors using the Neural Engineering Framework (see Appendix A for more details). Code for the model is available online at http://github.com/ikajic/spa-poem.

3.1. Structure of the model

Reviews of brain imaging studies of emotion have found that the relation between particular brain areas and specific emotions and emotional functions is highly complex (Kober et al., 2008; Lindquist et al., 2012). There is no one brain area that produces any particular emotion, and many brain areas are relevant to emotional functions such as physiological perception and cognitive appraisal. Accordingly, POEM’s neural networks and their connectivity are best interpreted functionally as operating across numerous brain areas including the amygdala, insula, prefrontal cortex, orbitofrontal cortex, and basal ganglia.

POEM’s structure is shown in Fig. 2. Inputs to the model are provided via the sensory and the interoception subnetworks by setting the desired vectors (e.g., “cat”, “snake” or “cake” for sensory, and “increased heartbeat“ for interoception). The sensory subnetwork models inputs from external senses such as eyes and ears, and the interoception subnetwork models inputs from internal physiological states such as heart rate. These inputs are then further processed in the memory-reasoning loop that includes subnetworks for episodic memory, language processing, and conceptualization, as suggested by Kober et al. (2008).

Within the memory-reasoning loop, the episodic subnetwork receives inputs from the sensory network and from the conceptualization subnetwork. The language subnetwork accesses the neural representations in episodic memory and uses them to combine neural representations corresponding to words, sending the results to the conceptualization subnetwork. Finally, the EPA subnetwork assigns to every concept represented in the episodic subnetwork a value for the evaluation, potency and activity dimensions of emotion. The executive subnetwork receives its input from the EPA subnetwork and generates a semantic pointer corresponding to a specific emotion label, based on locations of emotions in EPA space. The interactions of the subnetworks enable POEM to transform an input vector representing a situation into an output vector representing an emotional response to the situation.

Another important computational mechanism in POEM is the winner-take-all (WTA) algorithm implemented using lateral inhibition at the output of some subnetworks. In the *conceptualization* network, the WTA mechanism enables competition among semantic pointers. The competition is realized by mutual inhibition between groups of neurons and is determined by the strength of inhibition and synaptic timing parameters. The WTA is also used in the EPA subnetwork to clean up the noise in the semantic pointer representation of emotions. Those vectors that are below a certain threshold are inhibited, and those above the threshold facilitated.

3.2. Empirical data in the model

The EPA ratings for input concepts and output emotions are based on empirical studies that average participants’ responses along the dimensions of evaluation, potency, and activity. These ratings capture more than words, incorporating deep-meaning features of emotions such as immediate sensory perceptions and physiological reaction patterns included in emotions. Thus emotion components at varying levels of semantic depth of the EPA space serve as the common ground for integrating those different modalities in the identification of specific emotions resulting from various sensory inputs.

The empirically-derived datasets for EPA ratings employed in the simulations below were the International Affective Picture Set (Lang, Bradley, & Cuthbert, 2008) for sensory inputs, Fontaine et al.’s (2007) study of 144 emotion features for physiological and motor components of emotion, and a repository of words along with their affective EPA meanings compiled at Indiana University (Francis & Heise, 2006). Data collection for the Indiana repository is described in detail by Heise (2010). Simulation

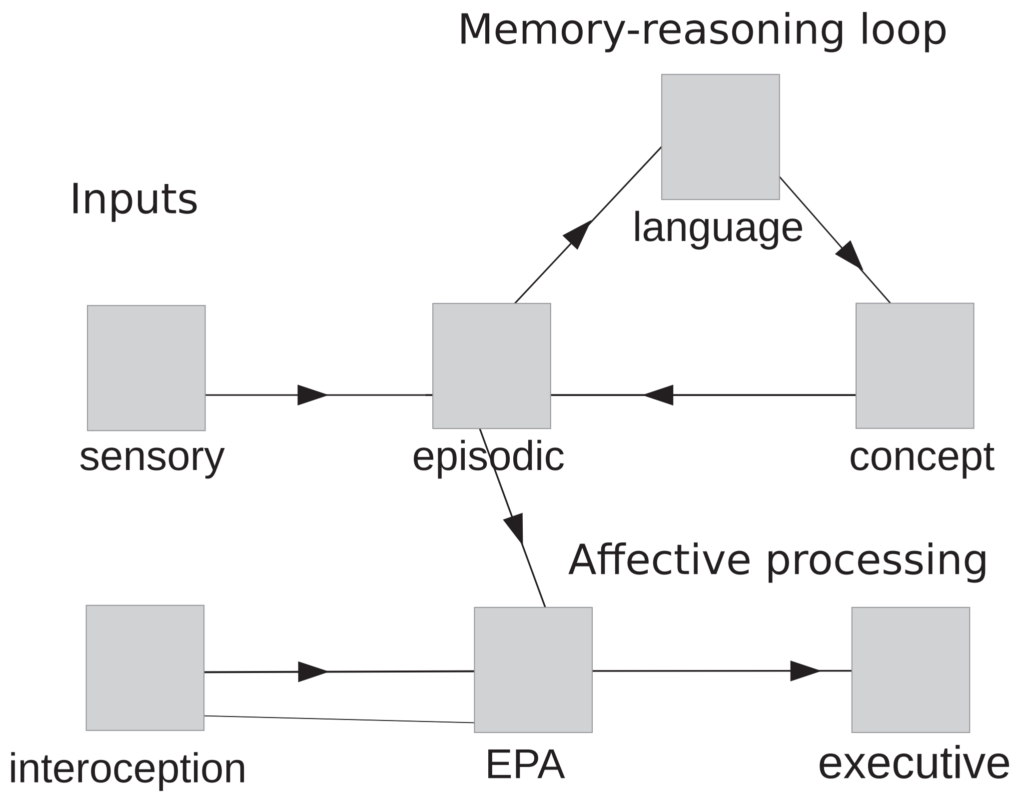

Fig. 2. The structure of the POEM model with rectangles as functionally distinct subnetworks of neurons. Arrowed lines indicate the flow of information between the subnetworks. Sensory and interoception subnetworks handle the inputs to the model, while episodic, language and conceptualization subnetworks make up the memory reasoning loop. The EPA and executive subnetworks assign an emotion label to the content in the episodic network.

5, which models reasoning about emotions, also employs coefficients from a statistical model of emotions arising from combinations of concepts rather than single words. The study and resulting statistical models are described by Heise (2010).

For emotion categorization, we use 23 emotions from Fontaine et al. (2007): anger, anxiety, being hurt, compassion, contempt, contentment, despair, disappointment, disgust, fear, guilt, happiness, hate, interest, irritation, jealousy, joy, love, pleasure, pride, sadness, shame and stress. Surprise was excluded from simulations due to concerns that it is not well captured by the three-dimensional EPA space (Fontaine et al., 2007). These emotions are implemented as firing patterns in the *executive* subnetwork of the POEM model. If no emotion is present, the affect network represents a vector with all zeros corresponding to the (0, 0, 0) EPA vector.

3.3. Simulations

POEM runs six simulations, which are listed in Table 1. These simulations cover a variety of emotion phenomena observed in everyday life and psychological experiments. The goal in choosing these simulations was to cover major theoretical perspectives on emotions reviewed more extensively below in the discussion section. Each simulation has stimuli represented as vectors presented to the sensory or interoception subnetwork and returns a pattern in the exec-

Table 1
Simulations and model inputs.

	Simulation	Sensory	Interoception
1.	Evolutionary roots of emotion	Snake	
2.	Dynamics of appraisal	Snake, glass, zoo	
3.	Embodiment		Smiled, frowned
4.	Interaction of physiological input and cognitive appraisal	Angry person, euphoric person	Felt hearbeat getting faster, muscles tensing whole body, felt breathing getting faster, sweated
5.	Reasoning about emotions	Mother, shout at, child, object, subject, action	
6.	Mixed emotions	Cake, taste, obesity, thought	

utive network as output corresponding to an emotion label. Note that the output label is strictly speaking not endogenous to the model as we hold the corresponding conscious emotion to be the pattern of neural firing activity. However, to allow the human modeler to evaluate the outcome of the simulation, we provide the best-matching verbal labels based on cosine similarity.

3.4. Results

Figs. 3–8 show the end state of each subnetwork given the inputs in Table 1, displaying a graph that shows the similarity between the ideal pattern of activity for a particular concept and the actual neural activity pattern. If the represented value matches the ideal one, the similarity value is close to one. Otherwise, if there is very little or no overlap between the patterns, the similarity value is close to zero. The simulation consists of presenting the input to the sensory and/or interoception networks, and reading out the corresponding emotional reactions in the executive network.

3.4.1. Simulation 1. Perception-based emotion

Some perceptual stimuli generate immediate emotional reactions. Simulation 1 simply presents a vector representation of a stimulus *snake*, based on data from Lang et al. (2008), as the input to the sensory network, resulting in the dynamics shown in Fig. 3. Setting the input triggers the semantic pointer representation of a concept *snake*, first in the sensory subnetwork, and then in the episodic, language and conceptualization subnetworks. In each of these networks the concept *snake* has a different neural representation, but each of them preserves the meaning of the concept. Each of the subnetworks in Fig. 3 shows the similarity between the vector at the input of the network and the stored semantic pointer. The affect network recognizes the input *snake* and assigns it an EPA value: $(-0.40, -0.50, 0.49)$. This is close to the empirical value $(-0.39, -0.48, 0.47)$, and the difference is a result of the noisy regime in which neurons representing numerical values operate. The EPA subnetwork connects to the executive subnetwork by mapping the EPA values to the closest semantic pointers for emotions. The resulting strongest emotions in the executive network are *anxiety, fear, shame* and *stress*, which all share strong negativity in the evaluation dimension in the EPA space. This models a quick pathway to the executive attention network that causes the fear-like emotions to become active. This may reflect an evolved tendency to become frightened immediately upon perceiving dangerous stimuli such as snakes and features of a predator's face (Cunningham, Van Bavel, & Johnsen, 2008).

3.4.2. Simulation 2. Dynamics of appraisal

According to appraisal theory, emotions follow from an ongoing process of cognitive evaluation of external stimuli with regard to the organism's goals. In the previous simulation modeling the experience of fear when seeing a snake, one could argue that fear results from the rapid assessment by the cognitive system that a predator endangers the sur-

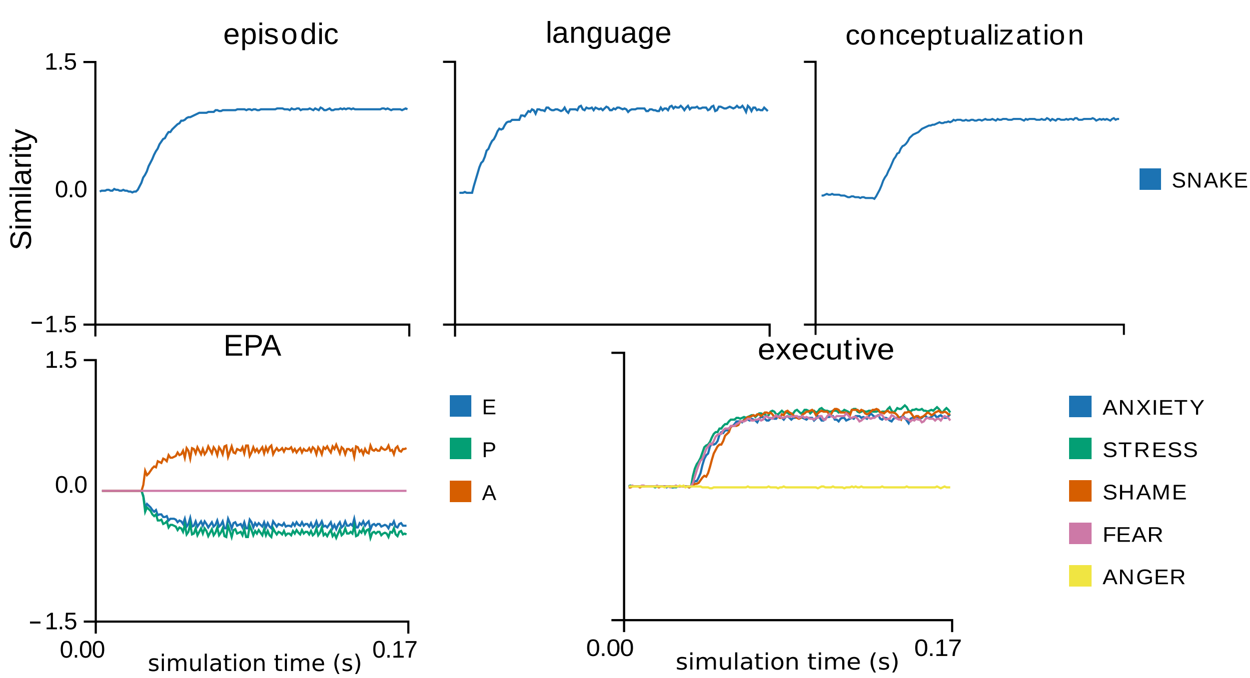

Fig. 3. Simulation 1. Perception-based emotions. The sensory subnetwork is presented with a vector for *snake*, which is recognized as a semantic pointer in the episodic subnetwork. This triggers the EPA subnetwork to produce the executive subnetwork response with negative emotions such as: *anxiety, fear* and *stress*. The Y axis shows the similarity, for concepts such as snake and anxiety, between the vector represented by the neurons in the subnetworks and the empirically-derived EPA vector for those concepts.

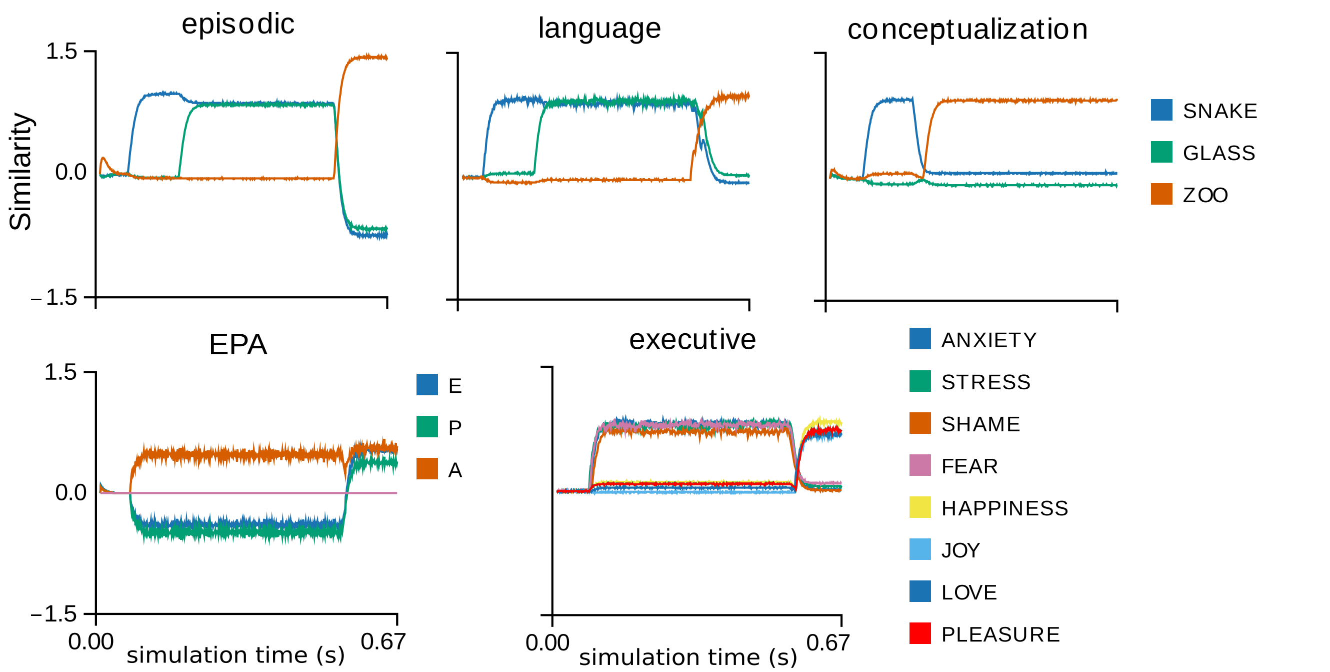

Fig. 4. Simulation 2. Dynamics of appraisal. The sensory network is first presented with the word snake (blue line), and then with words snake and glass (blue line and green line). After the last input has been presented, the representation of the concept *zoo* is triggered in the episodic network (orange line). This is evaluated to a set of positive emotions such as *happines, joy, love* and *pleasure*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

vival of the organism. However, cognitive appraisals of external events can also happen in much more deliberate, contextualized, and reflective ways, resulting in quite different emotional outcomes. For example, the emotion upon perceiving a snake might change from fear to interest if you are in a zoo and realize that you are safe by virtue of the glass keeping the snake in its terrarium, a process called re-appraisal (Scherer et al., 2001). POEM simulates the effects of cognitive re-appraisal on emotions in two steps (Fig. 4). First, as in the Simulation 1, the sensory network is presented with the input *snake* and triggers a similar set of emotions: *anxiety, stress, shame* and *fear* by means of a quick emotion pathway (Cunningham et al., 2008).

Second, we augment the input by activating semantic pointers *snake* and *glass*. The two semantic pointers then activate the semantic pointer *zoo*. Although the semantic pointer for *zoo* is somewhat similar to the semantic pointer *snake* in the conceptualization subnetwork, it will only get activated in the episodic network when both *snake* and *glass* are present. The delay in activation of the concept *zoo* is determined by the synaptic connections between populations of neurons. When *zoo* is conceptualized, the focus of attention from the sensory input is shifted to positive experiences related to the idea of a *zoo*, based on the EPA ratings (positive, powerful, and active) of the concept *zoo*. These ratings come from the Indiana study of affective meanings used as a database for our simulations (Francis & Heise, 2006). Positive experiences with *zoos* are manifested in a set of positive emotions in the executive network: *love, joy, happiness, pride* and *pleasure*. This slower emotion pathway invokes the language and conceptualization subnetworks and causes overriding of the initial negative emotional response by positive emotions. Thus, this simulation confirms that POEM is able to model cognitive reappraisal effects.

3.4.3. Simulation 3. Embodiment

Simulation 3 shows the effects of bodily states on the generation of emotion. Two inputs to the interoception network are used in two conditions, *smiled* and *frowned*, from the Fontaine et al. (2007) study. The resulting emotional reactions are shown in Fig. 5. A semantic pointer representing a *smile* is activated in the interoception network, resulting in emotions such as *happiness, pleasure, contentment* and *love*. Then, it is replaced by the semantic pointer *frowned*, which results in *hate, irritation* and *anger*. This behavior is consistent with experimental results showing both that emotions lead to reactions of associated facial muscles and that stimulation of facial muscles may lead to elicitation of the associated affective reactions (e.g., Niedenthal et al., 2009).

3.4.4. Simulation 4. Interaction of physiology and cognitive appraisal

Simulation 4 shows how the semantic pointer theory of emotion explains the experimental finding that specific emotional states emerge from the interaction of perceived physiological states and cognitive processes. The most famous demonstration of this interaction is an experiment

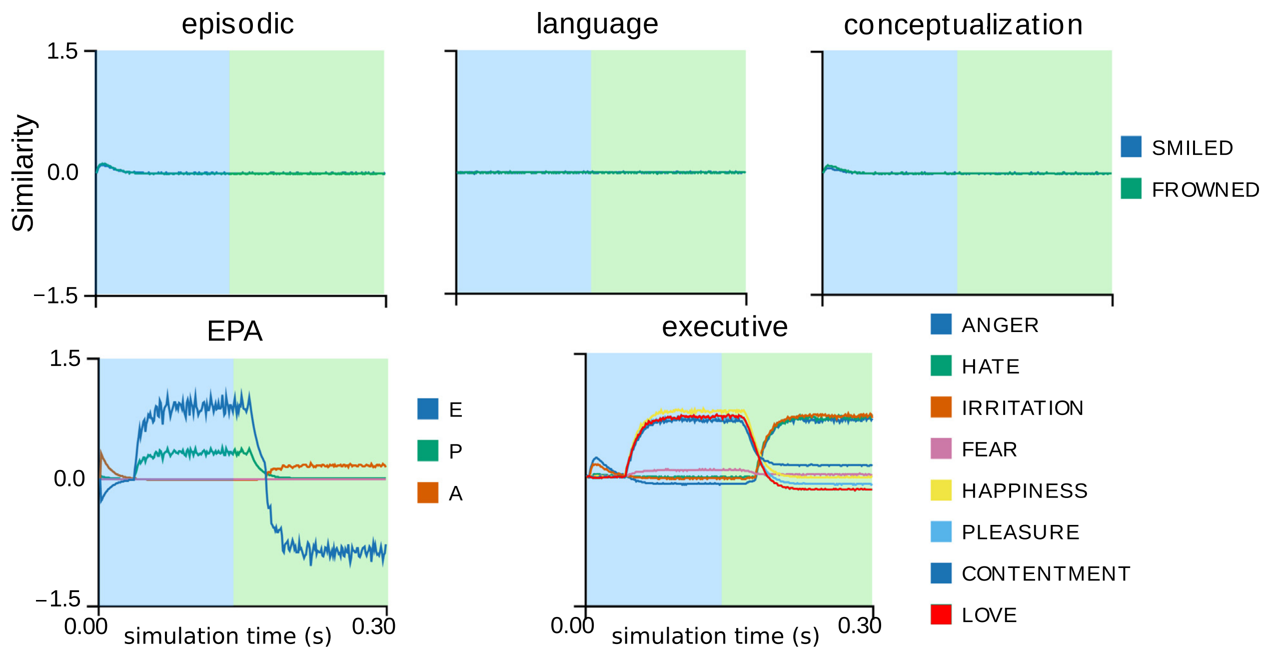

Fig. 5. Simulation 3. Embodiment. The interoception network is first presented with the input *smiled* (blue background) resulting in positive emotions in the executive network, such as *happiness, contentment* and *pleasure*. Then, the word *frowned* is presented (green background) resulting in negative emotions: *anger, irritation* and *hate*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where participants were transformed into states of euphoria or anger by manipulating the experimental context (Schachter & Singer, 1962). Participants were injected with adrenaline, which induced an increased heart rate, more rapid respiration, and muscle contractions. Then, they were asked to fill out questionnaire in two different contexts. The first context was pleasant, where a stooge was behaving euphorically. In the second context, the stooge appeared angry. The experiment showed that participants unaware of being injected with adrenaline attributed their arousal to feelings of euphoria in the condition with the euphoric stooge, and to anger in the condition with the angry stooge, in line with the claim that emotions require both a physiological base and a cognitive interpretation.

POEM simplifies the physiological effects of adrenaline by using only one semantic pointer, namely *felt heartbeat getting faster* as the input to the interoception network. Mapping of this semantic pointer to the EPA space is taken from the data reported by Fontaine et al. (2007). To simulate the effects on cognitive appraisals of the stooge's behavior as in the Schachter-Singer experiment, we also present either *angry* or *euphoric* to the sensory network. The behaviour of the model when sequentially presenting those inputs is shown in Fig. 6). When the input to the sensory network is set to semantic pointer *angry* it triggers negative emotions, mainly *anger* in the executive network. However, changing the input to *euphoric* triggers positive emotions such as *happiness, joy, pleasure* and *contentment*, closely simulating the euphoric state in the original experiment. This demonstrates that with the same physiological input, the communicative context is used to generate a different emotional state. To simulate the control condition of the experiment using POEM, we varied the physiological input and presented the vector *heartbeat slowing down*, while keeping the same sensory input. In this control condition, we observed lower potency and arousal values. If only using sensory input without simulating physiological arousal, the model thus still produces an emotional response, but is much more attenuated.

3.4.5. Simulation 5. Reasoning about emotions

Simulation 5 models the role of syntactic structure in emotion generation. Given a simple sentence: *The mother shouts at the child*, the model infers that both the child and the mother experience negative emotions. Emotions of the mother are *anger, hate* and *irritation*, while the child feels *shameful, anxious* and *fearful* (Fig. 7). The child also has some positive emotions such as *love* and *happiness* which arise through the association with the word *mother*.

The syntactic structure in the sentence is achieved by binding the semantic pointer *mother* with the semantic pointer *subject*, semantic pointer *shout at* with the semantic pointer *action* and the semantic pointer *child* with semantic pointer *object*. Every binding results in a new vector of the same dimensionality as the vectors which were bound together. Finally, all bindings are added together, which results in a compact vector-representation of a sentence. The semantic pointer *subject*, corresponding to the subject of the *shout at* action is assigned a value in the EPA subnetwork that has been computed with the INTERACT model

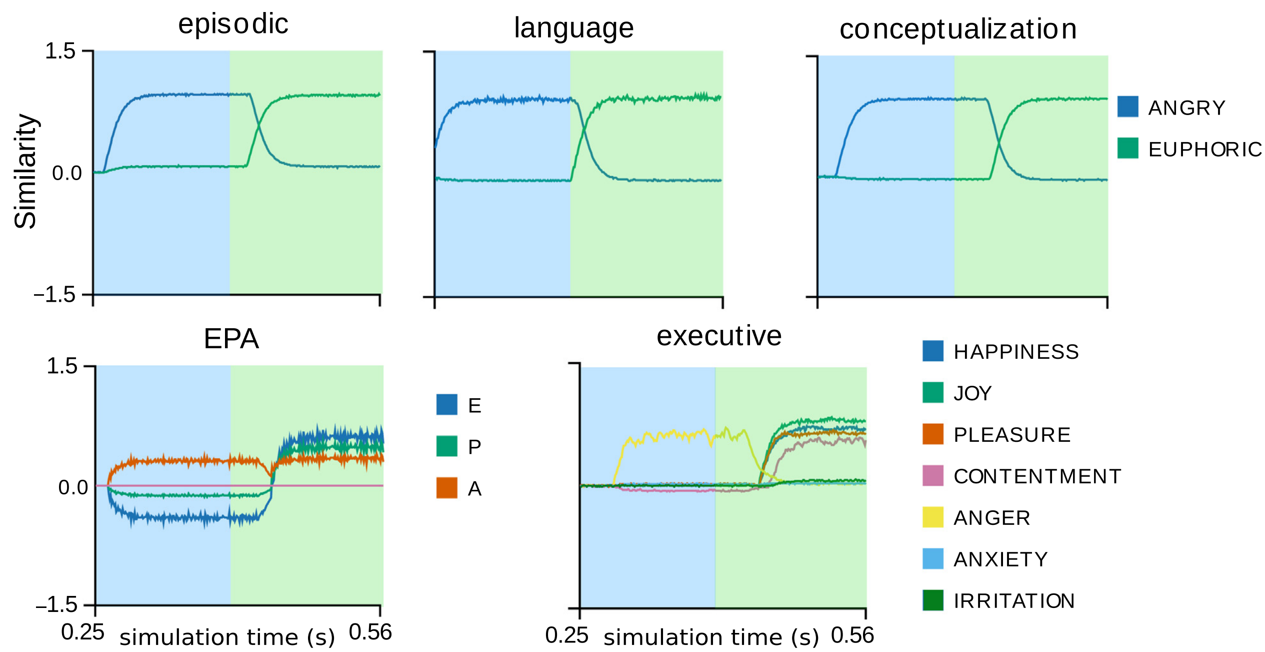

Fig. 6. Simulation 4. Interaction of physiological input and cognitive appraisal. Two different inputs are presented in a sequence to the sensory network simulating the mood of the stooge administering the experiment: angry and euphoric. The interoception network is presented with inputs simulating the effects of adrenaline injection. The model responds with negative emotions for the input *angry* and with positive emotions for the input *euphoric*.

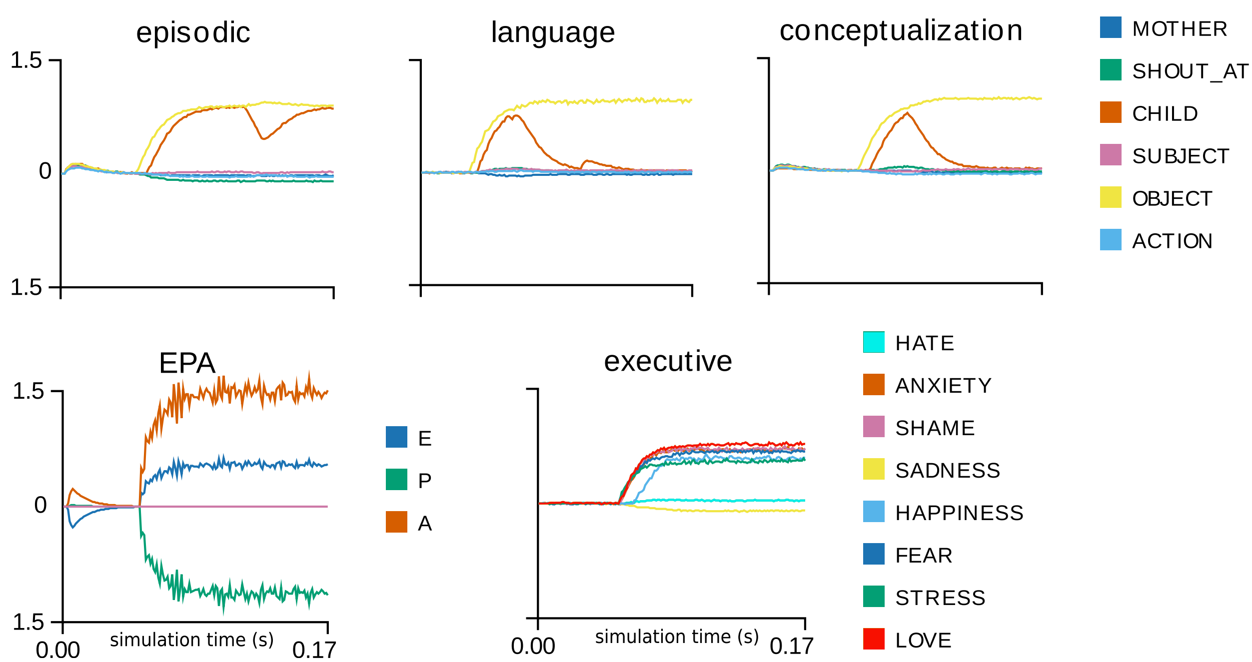

Fig. 7. Simulation 5. Reasoning about emotions. The model is presented with the sentence: *Mother shouts at child*. When queried about emotions of the child, the model responds with negative emotions such as *anxiety*, *fear*, *shame* and *stress*, but also with an activation of the positive emotions *love* and *happiness*, in line with the ambivalence of the mother-child relationship.

(Heise, 1997). The EPA values for the semantic pointer *object* are obtained in the same way. Semantic pointers for the *subject, object* and *action* are randomly generated. To find the emotions of a mother or the child, we use an additional component in the network that we label *query*. It takes the concepts *mother* or *child* as the input and allows extraction of the role (subject or object) given a person (mother or child). For more details about syntactic processing with semantic pointers see Appendix A.

3.4.6. Simulation 6. Mixed emotions

Simulation 6 demonstrates a scenario where a stimulus is appraised in multiple conflicting ways, to test whether POEM can also account for mixed emotions that result

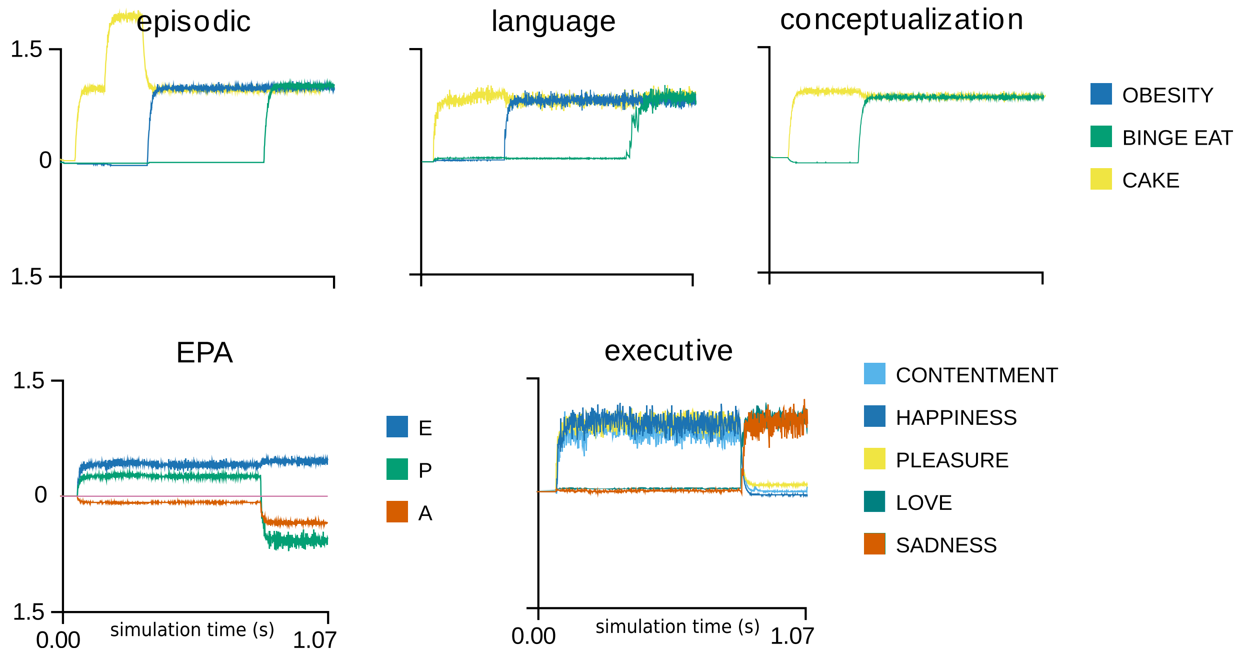

Fig. 8. Simulation 6. Mixed emotions. The model is first presented with the concept *cake*, which prompts positive emotions at the output. Then, the model is presented with the concept *obesity*, which triggers the representation of vector *binge eat*, finally resulting in mixed emotions *love* and *sadness*.

from ambivalent appraisals of events (cf., e.g., Larsen, McGraw, & Cacioppo, 2001). For example, eating a delicious cake evokes positive emotions such as joy. However, if one is on a diet, eating a piece of cake might also evoke negative emotions such as guilt. In this simulation we simulate this scenario using the binding of semantic pointers, introduced in Simulation 5. First, we present an input vector consisting of two bound vectors *cake* and *taste*. When we query for taste, we observe positive emotions such as *contentment, happiness* and *pleasure*. Then, we add additional input that consists of two bound vectors: *thought* and *obesity*, to simulate appraisals that might occur when eating a cake. This prompts a thought of binge eating in the episodic memory network, resulting in the change of the emotional response. Now, the emotions have changed to *love* and *sadness*, showing that PEOM can indeed explain mixed emotional experiences.

4. Discussion

We propose that emotions are semantic pointers, which are patterns of activity in neural populations that enable the brain to represent complex verbal and nonverbal information (Eliasmith, 2013). Semantic pointers can interrelate representations of bodily occurrences with symbolic cognitions through mechanisms of repeated binding. These properties of semantic pointers allow us to account for how emotions combine physiological changes in the body with cultural expectations embedded in linguistic structures. The implementation of the semantic pointer theory of emotions in a computational model contributes to the understanding of emotions as dynamical systems (cf. Scherer, 2009).

4.1. Strengths and limitations of the POEM model

POEM supports the semantic pointer theory of emotion by simulating important phenomena emphasized by different theoretical traditions in emotion research, covering embodiment, appraisal, and cultural construction. POEM models emotional phenomena from simple reflex-like emotion generation to reasoning about emotions in a social context, including mixed or ambivalent emotions. These simulations back our claim that the semantic pointer theory of emotion integrates physiological, cognitive, and social aspects of emotions.

There are two ways in which the semantic pointer theory of emotions could be falsified. The first would be the existence of some important aspect of emotions that the theory is incapable of explaining, for example if emotions could be experienced by disembodied souls. We have only simulated six phenomena, but they show an impressive range that could easily be extended. The second way of falsifying our theory would be to provide another theory with equal computational precision and greater explanatory power, modelling these phenomena and more. But to our knowledge, no current alternative incorporates physiology, cognition, and culture.

One limitation of the POEM model is that the neural subnetworks in the model do not correspond to anatomical structures in the human brain. Instead, the structure of POEM is inspired by appreciation that there is no one-to-one correspondence between emotions and brain areas (e.g., Kober et al., 2008; Lindquist et al., 2012). POEM provides a detailed account of neurocomputational mechanisms responsible for psychological functions required for emotions. Determining the exact anatomical structures that implement these functional computations is a matter for future empirical studies. When such information becomes available, the Neural Engineering Framework and the Nengo software that we used to produce POEM can be used for more anatomically detailed models (e.g., see Eliasmith et al., 2012). We hope to see in the future a more realistic version of POEM that respects the physical constraints of the computation of emotion found in the brain. It would be desirable, for example, to take into account the influence of neurotransmitters such as dopamine and serotonin, and hormones such as oxytocin and testosterone, on human emotion.

A further limitation is that the model with its current architectural constraints does not represent relevant cognitive functions such as perception, motivation, or behavioural control in realistic ways, but is restricted to rather semantic transformations. On the input side, one would have to explain how a complex sensory pattern evokes the conceptual representation of objects like snakes or cakes which start of the simulations described here. Note, however, that the Spaun model of the brain (Eliasmith et al., 2012) shows how such basic perceptive processes can be performed by the semantic-pointer architecture. On the output side, one would have to show how emotions are important contributors to action, encouraging people to perform or avoid specific behaviours that make them happy or fearful. Generating neural firing patterns corresponding to emotion labels is not enough to implement the functional properties of emotion in a cognitive architecture. But our POEM model meshes perfectly with semantic pointer accounts of automatic and intentional action developed in parallel (Schröder et al., 2014; Schröder & Thagard, 2013, 2014). Integrating existing semantic-pointer models of all the different cognitive processes from perception through emotion to action would be a daunting task not yet achievable. But we have contributed with the POEM model a demonstration how emotions can be understood in terms of semantic pointers, thus adding to the body of work supporting the claim the semantic-pointer architecture does provide a comprehensive and generic account of human cognition.

4.2. Relations with other theories of emotion

We now compare the semantic pointer theory of emotion with other major emotion theories, according to the four perspectives proposed by Gross and Barrett (2011). This discussion overlaps with the comparison of affect control theory (Heise, 2007) with the field of contemporary emotion theories provided by Rogers et al. (2014), as ideas from affect control theory regarding the conceptual and social base of emotions have been incorporated into the semantic pointer theory of emotions. However, the present theory goes beyond affect control theory by specifying neural mechanisms underlying emotion generation.

The semantic pointer theory of emotion is roughly compatible with the *psychological constructionist* perspective on emotions, providing a computational specification of mechanisms used by the brain for the generation of emotion (Thagard & Schröder, 2014). Psychological constructionism is the view that emotions are emergent products of more general mental operations that underlie other phenomena as well (Barrett, 2017; Barrett & Russell, 2014; Duncan & Barrett, 2007; Russell, 2009). We contribute to the constructionist project by using the Semantic Pointer Architecture, a general framework for linking psychological functions to biological computations performed by neurons. Our computational model POEM implements constructionist ideas precisely by simulating the generation of emotions dynamically through the interplay of functional neural subnetworks for perception, conceptualization, linguistic processing and so on, rather than through activation of local networks specific to particular emotion concepts.

A second perspective on emotions is *sociological*, emphasizing that emotions reflect shared conceptual structures in the cultural background of human thought and action (Heise, 2007; Hochschild, 1983; Kemper, 2006; Von Scheve, 2014). Many social constructionist approaches to emotion use qualitative methodologies that are difficult to unify with rigorous theories of emotion. Our implementation of the semantic pointer theory of emotions in the POEM model contributes to unification by using the techniques of affect control theory, treating emotions as mathematical transformations in a vector space constituted by the dimensions of evaluation, potency, and activity (Heise, 2007; Rogers et al., 2014). POEM provides a neuroscientific complement to affect control theory by showing how representations and transformations can be achieved in biologically plausible neural networks (see Schröder & Thagard, 2013, for a similar strategy).

This strategy accommodates cultural variation in emotional experience, (e.g., Mesquita & Frijda, 1992). Many studies based on affect control theory have looked at the affective basis of cross-cultural commonalities versus differences in human experience and action, including subcultural variation within increasingly heterogeneous societies (e.g., Ambrasat et al., 2014; Ambrasat, von Scheve, Schauenburg, Conrad, & Schröder, 2016; Heise, 2014; Schneider, 1996). Because the POEM model takes culture-specific affective connotations as input to simulating the generation of emotion, simulations can be sensitive to culture by using different empirically-based sentiment dictionaries (for empirical studies, see Heise, 2014; Moore, Romney, Hsia, & Rusch, 1999). Emotion-generating brains are sensitive to culture through the learned semantic relationships between specific patterns of neural activity.

The semantic pointer theory of emotion incorporates the *appraisal* perspective in two ways, only one of which is implemented in POEM. The conceptualization approach to emotion generation in POEM is different from the systematic evaluation of the relevance of events in relation to goals considered as the central mechanism of emotion generation in the different variants of appraisal theory (Oatley & Johnson-Laird, 1987; Ortony et al., 1990; Scherer et al., 2001). However, the dimensions of the EPA space employed for modeling emotion in affect control theory and in POEM is relevant to appraisal (Fontaine et al., 2007; Rogers et al., 2014; Scherer, Dan, & Flykt, 2006): The evaluation dimension corresponds to appraising the goal conduciveness of an object or event, potency to the person's perceived coping potential or control, and activity to the urgency of a behavioural reaction. One could thus argue that appraisal mechanisms are subtly built into the semantic structure of the representations that give rise to emotional experiences. From such a point of view, EPA vectors of concepts used in POEM do not represent intrinsic affective qualities of the concepts but rather reflect generic, de-contextualized semantic appraisals of these concepts by the raters that provided the EPA data in empirical studies.

Such semantic appraisal might partly result from biological evolution, as in simulation 1, where the pre-linguistic concept of a snake causes an evaluation-potency-activity pattern similar to the emotion of fear. Other situations require more recent cultural representations, such as the linguistic concept of a zoo in simulation 2, allowing the reaction to a snake to be overridden. Pre-linguistic concepts may have embedded appraisal patterns that reflect past experiences of the species, whereas linguistic concepts have appraisal patterns based on cultural learning. Either way, these concepts allow individuals to quickly rely on the knowledge of other individuals, significantly reducing the computational burden of an individual brain when it comes to appraising the goal relevance of a specific situation (for comparison of the computational tractability of appraisal theory versus affect control theory, see Hoey et al., 2016).

The second way of modelling appraisal compatible with the semantic pointer theory of emotion uses a neural network to model appraisal as a parallel process of identification of emotions based on parallel satisfaction of many goal-related constraints (Thagard & Aubie, 2008). In the future, it would be interesting to develop a model that combines the automatic, language-based appraisals now produced in POEM with the more systematic, computationally intensive goal-based appraisals performed by parallel constraint satisfaction.

The fourth major theoretical perspectives on emotions is the *basic emotions* approach, which treats emotions as relatively fixed, universal response programs embedded in the brain (Ekman & Cordaro, 2011). This view has been criticized as empirically untenable in light of recent neuroscientific advances (Lindquist et al., 2012). Accordingly, the semantic pointer theory of emotions and the POEM model may seem at odds with the basic emotion approach, but more contemporary basic-emotion approaches do not deny the cultural and social variability of emotional experiences, and allow that some aspects of emotional reactions and some classes of emotions are more inflexible and cross-culturally invariable than others (Ekman & Cordaro, 2011; Scarantino, 2014). This version of basic emotion theory may be reconcilable with semantic pointer theory, which allows for quick and possibly universal emotional reactions relatively unaltered by cultural influences, as in simulation 1 of a perceived snake quickly causing fear.

4.3. Outlook

Besides fitting with crucial tenets of major theories of emotion, the semantic pointer theory of emotion aligns well with related accounts of other cognitive phenomena. In our discussion of limitations of the POEM model, we already pointed out the close similarities to existing semantic-pointer models of both automatic and intentional action (Schröder et al., 2014; Schröder & Thagard, 2013, 2014). Similarly, the semantic pointer architecture has been employed to explain a variety of cognitive phenomena in general such as visual perception, categorization, memory, and motor control (Blouw et al., 2016; Eliasmith, 2013; Eliasmith et al., 2012). While it is not feasible (and perhaps not even desirable) to create a unified model of all neurocognitive processes, our POEM model certainly adds to the number of phenomena that can be explained with the semantic pointer idea, thus enhancing both a "unified mind" approach to emotion theory (cf. Barrett & Russell, 2014) and the status of semantic pointers as a general cognitive architecture (cf. Eliasmith, 2013).

Happiness, sadness, pride, shame, and hundreds of other emotions all involve conscious feelings that cannot be ignored in a comprehensive theory of emotions. Fortunately, the theory of consciousness as semantic pointer competition employs many of the same mechanisms as our account of emotion (Thagard & Stewart, 2014). Representation by neural firing, binding of representations into more complex semantic pointers, and competition among semantic pointers generate conscious experiences. Emotions differ from other conscious experiences such as perceptions and thoughts because of different inputs to neural representations and because of different bindings that incorporate both physiology and cognitive appraisal.

Future research should address empirical tests of the semantic pointer theory of emotion beyond its capability of simulating six emotional phenomena. The advantage of a detailed computational model with empirically-based inputs to simulations is a high amount of predictive precision compared to other theories of emotion limited to verbal descriptions. Improved understanding of emotions should come from triangulating model predictions with location information obtained from brain imaging studies and temporal information obtained from EEG studies (e.g., Conrad, Recio, & Jacobs, 2011; Schauenburg et al., 2019). For example, new brain scanning techniques such as diffusion tensor imaging are making it possible to investigate dynamic interactions among different brain areas. We predict that such techniques will reveal that emotional brains have rich interactions among cortical areas such as the orbitofrontal cortex relevant to appraisal and subcortical areas such as the amygdala relevant to physiological perception, with corresponding activation in association areas where binding of semantic pointers can occur. Because the Semantic Pointer Architecture is now implemented on neuromorphic chips that control simple robots, future simulations could combine appraisal and bodily perception in a physical robot. Neuromorphic hardware also may provide energy-efficient virtual agents that perform real-time sentiment analysis using the techniques of emotion coding that we have described.

Possible applications of the semantic pointer theory of emotions include clinical phenomena. All mental disorders have substantial emotional components, for example the excessive sadness that constitutes depression and the intense fear that accompanies paranoid schizophrenia. A theory of emotion should have implications both for explaining such disorders and for suggesting effective treatments. Thagard (2019b) argues that depression can be explained as the result of breakdowns in semantic pointer mechanisms, which are also relevant to understanding the benefits of treatments using therapy (i.e., altered linguistic appraisal patterns as in simulation 2) and antidepressants (i.e., altered physiological inputs as in simulations 3 and 4).

Political ideology is another important application of emotions (cf. Westen, 2008). Homer-Dixon et al. (2013) have proposed a theory of ideological dynamics that emphasizes the role of concepts tied to emotions. The semantic pointer approach provides precise mechanisms of emotion, thought, and action relevant to politics and many other social phenomena. Dealing with social problems requires integrating physiological, cognitive, neural, and social aspects of emotions.

Appendix A. Neural modeling[1]

To construct the computational model shown in this paper, we make use of the Neural Engineering Framework (NEF; Eliasmith et al., 2003). In this approach, we specify a type of distributed representation for each group of neurons, and we analytically solve for the connection weights between neurons that produce the desired computations. The distributed representation is a pattern of firing activity of a group of neurons. A group of neurons can represent various stimuli such as a word (e.g. *offer* or *smoke*), an image or a sound. Any stimulus which can be expressed as a vector of numerical values can be stored as a multidimensional pattern. As a result, every concept is represented as a random unique vector. All vectors have been normalized to length one for computational convenience. To leverage the ability of neurons to represent concepts, we must define how a group of neurons can store a vector using spiking activity (*encoding*). To make sense of such vector, we also need to define methods that assign a concept to each represented vector (*decoding*).

To define this neural encoding, the NEF generalizes standard results from sensory and motor cortices (Georgopoulos, Schwartz, & Kettner, 1986) that in order to represent a vector, each neuron in a population has a random "preferred direction vector" – a particular vector for which that neuron fires most strongly. This preference is biologically defined by the properties of neurons which depend on the location of the neuron within a network of neurons and its intrinsic characteristics. The more different the current vector is from that preferred vector, the less quickly the neuron will fire. In particular, Eq. (1) gives the amount of current J that should enter a neuron, given a represented vector $\boldsymbol{x}$, a preferred direction vector $\boldsymbol{e}$, a neuron gain α and a background current J_{bias}. The parameters α and J_{bias} are randomly chosen from a uniform distribution. Adjusting their statistical distribution produces neurons that give realistic background firing rates and maximum firing rates (Fig. 4.3 in Eliasmith et al., 2003). These parameters also impact the model itself; for example, having an overall lower average firing rate means that the model will require more neurons to produce the same level of accuracy.

$$J = \alpha \boldsymbol{e}\boldsymbol{x} + J_{bias} \tag{1}$$

This current can then be provided as input to any existing model of an individual neuron. The neuron response to input vector $\boldsymbol{x}$ is a specific spike pattern. For this paper, we used the standard Leaky Integrate-and-Fire neuron model, which is a simple model that captures the behaviour of a wide variety of observed neurons (Koch, 2004). Input current causes the membrane voltage V to increase as per Eq. (2), with neuron membrane resistance R and time constant τ_{RC}. For the models presented here, τ_{RC} was fixed at 20 ms (Isokawa, Levesque, Fried, & Engel, 1997). When the voltage reaches a certain threshold, the neuron fires (emits a spike), and then resets its membrane voltage for a fixed refractory period. For simplicity, we normalize the voltage range such that the reset voltage to 0, the firing threshold is 1, and R is also 1. In the absence of input, the voltage decays to the resting state value due to leak currents.

$$\frac{dV}{dt} = \frac{JR - V}{\tau_{RC}} \tag{2}$$

Given Eqs. (1) and (2), we can covert any vector $\boldsymbol{x}$ into a spiking pattern across a group of realistically heterogenous neurons. Furthermore, we can use Eqs. (3) and (4) to convert that spiking pattern back into an estimate of the original $\boldsymbol{x}$ value. This lets us determine how accurately the neurons are representing given values. Higher accuracy can be achieved by employing more neurons, at the computational cost of memory and time. The intuitive idea behind Eq. (3) is that we can take the average activity a of each neuron i, and estimate the represented value $\boldsymbol{x}$ by finding a fixed weighting factor $\boldsymbol{d}$ for each neuron. Eq. (4) shows how to solve for the optimal $\boldsymbol{d}$ as a least-squared error min-

[1] This appendix is largely identical with the appendix to Schröder et al. (2014), with modifications related to specific details of the present POEM model.

imization problem, where the sum is over a random sampling of the possible $\boldsymbol{x}$ values.

$$\hat{\boldsymbol{x}} = \sum_i a_i d_i \tag{3}$$

$$d_i = \sum_j \Gamma_{ij}^{-1} \Upsilon_j \quad \Upsilon_i = \sum_x a_i x \quad \Gamma_{ij} = \sum_x a_i a_j \tag{4}$$

Eqs. (3) and (4) allow the interpretation of the spiking data coming from the models. We take the spiking output of each neuron, decode it to an estimate of $\boldsymbol{x}$ and compare that to the ideal vectors for the various concepts in the model. If these vectors are close, then we add the text labels (e.g. *mother*, *love*, *child*) to the graphs, indicating that the pattern is very similar to the expected pattern for those terms.

It should be noted that this produces a generic method for extracting $\boldsymbol{x}$ from a spiking pattern without requiring a specific set of $\boldsymbol{x}$ values to optimize over. That is, we can accurately use $\boldsymbol{d}$ to determine if a particular pattern of activity means *mother* even though we do not use the *mother* vector to compute $\boldsymbol{d}$. The sums used to compute $\boldsymbol{d}$ in Eq. (4) are over a random sampling of $\boldsymbol{x}$. Since $\boldsymbol{x}$ in our simulations covers a 512-dimensional vector space and because we use less than 100 samples in that space, it is highly unlikely that the sampling includes exactly the vector for *mother* (or any other semantic pointer), but as shown in simulation plots, we can still use $\boldsymbol{d}$ to identify the presence of those semantic pointers (or any others).

Importantly, we also use Eq. (4) to compute the connection weights between groups of neurons. In contrast to other neural modelling methods which rely on learning, the NEF optionally allows us to directly compute connection weights that will cause neural models to behave in certain ways. For example, given two groups of neurons, we can form connections between them that will pass whatever vector is represented by one group to the next group by using the connection weights given in Eq. (5) (see Eliasmith et al., 2003, for detailed proof).

$$\omega_{ij} = \alpha_j \boldsymbol{e_j} \cdot \boldsymbol{d_i} \tag{5}$$

The weights allow us to pass information from one group to another, but to implement transition rules we need to compute non-linear transformations. In NEF, this can be achieved by computing $\boldsymbol{d}$ values for arbitrary functions $\boldsymbol{f}(\boldsymbol{x})$:

$$d_i^f = \sum_j \Gamma_{ij}^{-1} \Upsilon_j \quad \Upsilon_i = \sum_x a_i \boldsymbol{f}(\boldsymbol{x}) \quad \Gamma_{ij} = \sum_x a_i a_j \tag{6}$$

As a result, if the first neural population represents a vector $\boldsymbol{x}$, the connections will cause the second population to fire a pattern representing $\boldsymbol{f}(\boldsymbol{x})$. This approach allows us to compute transition rules which map the particular input vectors to particular output vectors in each of the neural networks in the model.

In the last simulation, we use the operation known as *circular convolution* to extract the subject or the object of the action in a given sentence. Circular convolution takes two vectors ($\boldsymbol{x}$ and $\boldsymbol{y}$) and produces a third vector $\boldsymbol{z}$, as per Eq. (7). The vector $\boldsymbol{z}$ can be thought of as a compressed representation of $\boldsymbol{x}$ and $\boldsymbol{y}$, forming the basis of our semantic pointers. Importantly, given $\boldsymbol{z}$ and $\boldsymbol{y}$ (or $\boldsymbol{x}$) we can recover an approximation of $\boldsymbol{x}$ (or $\boldsymbol{y}$) by computing the circular correlation (the inverse operation of circular convolution) as per Eq. (8).

$$z_i = \sum_j \boldsymbol{x}_j \boldsymbol{y}_{i-j} \tag{7}$$

$$\hat{\boldsymbol{x}}_i = \sum_j \boldsymbol{z}_j \boldsymbol{y}_{i+j} \tag{8}$$

This is how semantic pointers can be decompressed in their constituents. We use this property in Simulation 5, where the sensory input is a sentence which contains a combination of role and filler words. Roles are *subject*, *object* and *action*, and filler words are *mother*, *child* and *shout at*, respectively. To ensure that each filler word is assigned to the right role, we apply circular convolution to role-filler pairs and sum the resulting vectors. To retrieve individual roles, we can apply circular correlation to the sentence vector. By correlating the vector *child* with the above sentence, we obtain the vector *object*. The resulting vector will not be exactly the vector for *object* due to small amounts of noise resulting from correlation with other word pairs (e.g. *mother*∗*subject* and *action*∗*shout at*). The noise is there as vectors are not perfectly orthogonal, but they are sufficiently dissimilar so the noise can be cleaned up from the output. It is important to notice that this syntactic decomposition of role-filler pairs would not have been possible if we simply added the vectors: *subject* + *object* + *action* + *mother* + *child* + *shout at*. Given such a vector, it would not be possible to determine which filler word was assigned to which role word.

Appendix B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.cogsys.2019.04.007.

References

Ambrasat, J., von Scheve, C., Conrad, M., Schauenburg, G., & Schröder, T. (2014). Consensus and stratification in the affective meaning of human sociality. *Proceedings of the National Academy of Sciences of the United States of America, 111*(22), 8001–8006.

Ambrasat, J., von Scheve, C., Schauenburg, G., Conrad, M., & Schröder, T. (2016). Unpacking the habitus: Meaning making across lifestyles. *Sociological Forum, 31*(4), 994–1017.

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.

Balkenius, C., CaAsamero, L., PAdrnamets, P., Johansson, B., Butz, M. V., & Olsson, A. (2016). Outline of a sensory-motor perspective on intrinsically moral agents. *Adaptive Behavior, 24*(5), 306–319.

Barrett, L. F. (2009). The future of psychology: Connecting mind to brain. *Perspectives on Psychological Science, 4*(4), 326–339.

Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.

Barrett, L. F., Lewis, M., & Haviland-Jones, J. M. (2016). *Handbook of emotions*. Guilford.

Barrett, L. F., & Russell, J. A. (2014). *The psychological construction of emotion*. Guilford Publications.

Barsalou, L. (2016). On staying grounded and avoiding quixotic dead ends. *Psychonomic Bulletin and Review, 23*, 1122–1142.

Belli, S., Harré, R., & Iñiguez-Rueda, L. (2010). What is love? Discourse about emotions in social sciences. *Human Affairs, 20*(3), 249–270.

Blouw, P., Solodkin, E., Thagard, P., & Eliasmith, C. (2016). Concepts as semantic pointers: A framework and computational model. *Cognitive Science, 40*, 1128–1162. https://doi.org/10.1111/cogs.12265.

Conrad, M., Recio, C., & Jacobs, A. M. (2011). The time course of emotion effects in first and second language processing: A cross cultural erp study with german-spanish bilinguals. *Frontiers in Psychology, 2*.

Craig, A. D. (2002). How do you feel? Interoception: The sense of the physiological condition of the body. *Nature Reviews Neuroscience, 3*(8), 655–666.

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

Crawford, E., Gingerich, M., & Eliasmith, C. (2016). A functional architecture of the human brain: Emerging insights from the science of emotion. *Cognitive Science, 40*, 782–821.

Cunningham, W. A., Van Bavel, J. J., & Johnsen, I. R. (2008). Affective flexibility evaluative processing goals shape amygdala activity. *Psychological Science, 19*(2), 152–160.

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

Damasio, A. R., & Carvalho, G. B. (2013). The nature of feelings: Evolutionary and neurobiological origins. *Nature Reviews Neuroscience, 14*(2), 143–152.

Duncan, S., & Barrett, L. F. (2007). Affect is a form of cognition: A neurobiological analysis. *Cognition and Emotion, 21*(6), 1184–1211.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review, 3*(4), 364–370.

Eliasmith, C. (2013). *How to build a brain: A neural architecture for biological cognition*. New York, NY: 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., & Rasmussen, D. (2012). A large-scale model of the functioning brain. *Science, 338*, 1202–1205. https://doi.org/10.1126/science.1225266.

Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General, 113*(3), 464–486.

Feldman, J. (2013). The neural binding problem (s). *Cognitive Neurodynamics, 7*(1), 1–11.

Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science, 18*(12), 1050–1057.

Francis, C., & Heise, D. R. (2006). Mean affective ratings of 1500 concepts by Indiana university undergraduates in 2002–3. *Data in Computer Program Interact*.

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science, 233*(4771), 1416–1419.

Gratch, J., Cheng, L., & Marsella, S. (2015). The appraisal equivalence hypothesis: Verifying the domain-independence of a computational model of emotion dynamics. In *In 2015 international conference on affective computing and intelligent interaction (ACII)* (pp. 105–111). IEEE.

Gratch, J., & Marsella, S. (2004). A domain-independent framework for modeling emotion. *Cognitive Systems Research, 5*(4), 269–306.

Gross, J. J., & Barrett, L. F. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review, 3*(1), 8–16.

Heise, D. R. (1997). *INTERACT: Introduction and software. Affect Control Theory website*. University of Indiana, Retrieved from ~socpsy/ACT/interact.htm" xlink:type="simple">http://www.indiana.edu/~socpsy/ACT/interact.htm.

Heise, D. R. (2007). *Expressive order: Confirming sentiments in social actions*. Springer.

Heise, D. R. (2010). *Surveying cultures: Discovering shared conceptions and sentiments*. Hoboken: Wiley.

Heise, D. R. (2014). Cultural variations in sentiments. *SpringerPlus, 3*(1), 1.

Hochschild, A. R. (1983). *The managed heart: Commercialization of human feeling*. Berkeley: University of California Press.

Hoey, J., Schröder, T., & Alhothali, A. (2016). Affect control processes: Intelligent affective interaction using a partially observable markov decision process. *Artificial Intelligence, 230*, 134–172.

Homer-Dixon, T., Maynard, J. L., Mildenberger, M., Milkoreit, M., Mock, S. J., Quilley, S., & Thagard, P. (2013). A complex systems approach to the study of ideology: Cognitive-affective structures and the dynamics of belief systems. *Journal of Social and Political Psychology, 1*(1), 337–363.

Hudlicka, E. (2011). Guidelines for designing computational models of emotions. *International Journal of Synthetic Emotions (USE), 2*(1), 26–79.

Isokawa, M., Levesque, M., Fried, I., & Engel, J. (1997). Glutamate currents in morphologically identified human dentate granule cells in temporal lobe epilepsy. *Journal of Neurophysiology, 77*(6), 3355–3369.

James, W. (1884). What is an emotion? *Mind, 8*(34), 188–205.

Kemper, T. D. (2006). *Power and status and the power-status theory of emotions. In Handbook of the sociology of emotions*. Springer (pp. 87–113). Springer.

Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., & Wager, T. D. (2008). Functional grouping and cortical-subcortical interactions in emotion: A meta-analysis of neuroimaging studies. *Neuroimage, 42*(2), 998–1031.

Koch, C. (2004). *Biophysics of computation: Information processing in single neurons*. Oxford University Press.

Lakoff, C., & Johnson, M. (2003). *Metaphors we live by*. University of Chicago Press.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International affective picture system (iaps): Affective ratings of pictures and instruction manual. *Technical Report*, A-8.

Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology, 81*(4), 684.

Lindquist, K. A., & Barrett, L. F. (2012). A functional architecture of the human brain: Emerging insights from the science of emotion. *Trends in Cognitive Sciences, 16*(11), 533–540.

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*(03), 121–143.

Lively, K. J., & Heise, D. R. (2004). Sociological realms of emotional experiencel. *American Journal of Sociology, 109*(5), 1109–1136.

Lively, K. J., & Heise, D. R. (2014). Emotions in affect control theory. *Handbook of the sociology of emotions* (Vol. 2, pp. 51–75). Springer.

Marsella, S., & Gratch, J. (2014). Computationally modeling human emotion. *Communications of the ACM, 57*(12), 56–67.

Marsella, S., & Gratch, J. (2016). Computational models of emotion as psychological tools. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 113–129). Guilford.

Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions: A review. *Psychological Bulletin, 112*(2), 179.

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*(3), 529–546.

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

Oatley, K. (2009). Communications to self and others: Emotional experience and its skills. *Emotion Review, 1*(3), 206–213.

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion, 1*(1), 29–50.

Ortony, A., Clore, G. L., & Collins, A. (1990). *The cognitive structure of emotions*. Cambridge University Press.

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

Pessoa, L. (2013). *The cognitive-emotional brain: From interactions to integration*. MIT Press.

Picard, R. W. (1997). *Affective computing*. Cambridge: MIT Press.

Rasmussen, D., Voelker, A., & Eliasmith, C. (2017). A neural model of hierarchical reinforcement learning. *PloS One, 12*(7), e0180234.

Reisenzein, R., Hudlicka, E., Dastani, M., Gratch, J., Hindriks, K., Lorini, E., & Meyer, J.-J. C. (2013). Computational modeling of emotion: Toward improving the inter- and intradisciplinary exchange. *IEEE Transactions on Affective Computing, 4*(3), 246–266.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.

Rogers, K. B., Schröder, T., & von Scheve, C. (2014). Dissecting the sociality of emotion: A multilevel approach. *Emotion Review, 6*(2), 124–133.

Rumelhart, D. E., McClelland, J. L., Group, P. R., et al. (1986). Parallel distributed processing: Explorations in the microstructure of cognition, Vols. 1–2. Cambridge, MA.

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition and Emotion, 23*(7), 1259–1283.

Scarantino, A. (2014). Basic emotions, psychological construction, and the problem of variability. In L. F. Barrett & J. A. Russel (Eds.), *The psychological construction of emotion* (pp. 334–376). Guilford.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69*(5), 379.

Schauenburg, G., Conrad, M., von Scheve, C., Barber, H., Ambrasat, J., Aryani, A., & Schröder, T. (2019). Making sense of social interaction: Emotional coherence drives semantic integration in the brain as assessed by event-related potentials. *Neuropsychologia, 125*, 1–13.

Scherer, K. R. (2009). Emotions are emergent processes: They require a dynamic computational architecture. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1535), 3459–3474.

Scherer, K. R., BAdnziger, T., & Roesch, E. (2010). *A blueprint for affective computing: A sourcebook and manual*. Oxford University Press.

Scherer, K. R., Dan, E., & Flykt, A. (2006). What determines a feeling's position in affective space? A case for appraisal. *Cognition & Emotion, 20*(1), 92–113.

Scherer, K. R., Schorr, A., & Johnstone, T. (2001). *Appraisal processes in emotion: Theory, methods, research. Series in Affective Science*. USA: Oxford University Press.

Schneider, A. (1996). Sexual-erotic emotions in the us in cross-cultural comparison. *International Journal of Sociology and Social Policy, 16*(9/10), 123–143.

Schröder, T., Hoey, J., & Rogers, K. B. (2016). Modeling dynamic identities and uncertainty in social interaction: Bayesian affect control theory. *American Sociological Review, 81*, 828–855.

Schröder, T., Stewart, T. C., & Thagard, P. (2014). Intention, emotion, and action: A neural theory based on semantic pointers. *Cognitive Science, 38*(5), 851–880.

Schröder, T., & Thagard, P. (2013). The affective meanings of automatic social behaviors: Three mechanisms that explain priming. *Psychological Review, 120*(1), 255–280.

Schröder, T., & Thagard, P. (2014). Priming: Constraint satisfaction and interactive competition. *Social Cognition, 32*, 152–167.

Singer, W. (2007). Binding by synchrony. *Scholarpedia, 2*(12), 1657. https://doi.org/10.4249/scholarpedia.l657, revision #124403.

Stewart, T. C., Tripp, B., & Eliasmith, C. (2009). Python scripting in the nengo simulator. *Frontiers in Neuroinformatics, 3*. https://doi.org/10.3389/neuro.11.007.2009.

Thagard, P. (2012a). *The brain and the meaning of life*. Princeton: Princeton University Press.

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

Thagard, P. (2019a). *Brain-mind: From neurons to consciousness and creativity*. Oxford: Oxford University Press.

Thagard, P. (2019b). *Mind-society: From brains to social sciences and professions*. Oxford: Oxford University Press.

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*(3), 811–834.

Thagard, P., & Schröder, T. (2014). Emotions as semantic pointers: Constructive neural mechanisms. In L. F. Barrett & J. A. Russel (Eds.), *The psychological construction of emotion* (pp. 144–167). Guilford.

Thagard, P., & Stewart, T. C. (2014). Two theories of consciousness: Semantic pointer competition vs. information integration. *Consciousness and Cognition, 30*, 73–90.

Von Scheve, C. (2014). *Emotion and social structures: The affective foundations of social order*. New York: Routledge.

Westen, D. (2008). *The political brain: The role of emotion in deciding the fate of the nation*. PublicAffairs.