ARTICLE

emotion review

Emotion Review
Vol. 13, No. 3 (July 2021) 245–256

ISSN 1754-0739
DOI: 10.1177/17540739211014946
https://journals.sagepub.com/home/emr

Operationalizing the Relation Between Affect and Cognition With the Somatic Transform

Neil J. MacKinnon
Department of Sociology & Anthropology, University of Guelph, Canada

Jesse Hoey
Cheriton School of Computer Science, University of Waterloo, Canada

Abstract

This article introduces the somatic transform that operationalizes the relation between affect and cognition at the psychological level of analysis by capitalizing on the relation between the cognitive-denotative and affective-connotative meaning of concepts as measured with semantic differential rating scales. Following discussion of levels of analysis, the importance of language at the psychological level, and two principles (inextricability and complementarity) summarizing the relation between affect and cognition that are rendered explicit by the somatic transform, we present affect control theory (ACT) and its Bayesian extension (BayesACT) containing the somatic transform. We conclude by identifying examples of inextricability and complementarity in the social science and neuroscience literatures and discussing how our psychological model might be implemented in a realistic neural model.

Keywords
affect, affect control theory, affective-connotative meaning, Bayesian affect control theory, cognition, cognitive-denotative meaning, emotion

The relation between cognition and affect has been a long-standing and often contentious issue in both the biological and social sciences, as exemplified by special issues of journals devoted to the topic, *Cognition and Emotion* in 2007 (Eder et al., 2007a) and *Current Opinion in Behavioral Sciences* in 2018 (Mather & Fanselow, 2018). Among other things, the closely related distinction between rational and emotional thinking has engendered a number of dual-process theories in the social sciences.[1] The affect–cognition divide has also contributed to fragmentation within academic disciplines, exemplified in the social sciences by the coexistence of cognitive psychology and the psychology of emotions,[2] and of cognitive sociology and the sociology of emotions.[3] In social psychology, the "cognitive–affective crossfire" has been used to describe the debate over the relative motivational power of cognitive self-consistency versus affective self-enhancement (MacKinnon, 2015; Swann et al., 1987), as well as the relative merits of identity theories emphasizing cognitive processes (Burke & Reitzes, 1991; Burke & Stets, 2009; Stryker, 1980) and those emphasizing affective processes (Heise, 1979, 2007; MacKinnon, 1994.) The cognition–affect divide also distinguishes transactional exchange theory, predicated on cognitive-rational assumptions about human motivation and social behavior, from relational exchange theory based on affective ties and commitment (Lawler et al., 2015, 2009).

Fragmentation along the affect–cognition divide has also occurred in neuroscience. Because of its traditional focus on cognitive processes, affective neuroscience (Damasio, 1994; LeDoux, 1996; Panksepp, 1998) emerged as a largely separate development; see the special issue on the emergence of affective neuroscience of *Brain and Cognition* (Schmidt, 2003).

Like neuroscience, artificial intelligence (AI) began with an emphasis on cognition, conceptualizing the human mind as a cognitive, information-processing system, and treating emotion as epiphenomenal noise or, at best, a functional subordinate of cognition (Johnson-Laird & Oatley, 1992; Simon, 1967). As a consequence, AI has failed to create an adequate model of the human mind, something many believe is essential to guide the successful mapping of perception and information-processing to motivated action. Over time, however, this narrow cognitive

Corresponding author: Neil J. MacKinnon, Department of Sociology & Anthropology, University of Guelph, Guelph, Ontario, N1G2, Canada.
Email: nmackinn@uoguelph.ca

view became challenged with the advent of emotional AI or affective computing (Picard, 1997).

Despite the historical prominence and sometimes deleterious effects of the affect–cognition divide for the study of emotion, recent decades have witnessed something of a ceasefire. While maintaining some degree of distinction, most emotion scholars and researchers today accept that affect and cognition are overlapping and functionally related processes of the human mind. What is lacking, however, are more precise ways to operationalize this relationship.

This article introduces the somatic transform[4] that operationalizes the relation between affect and cognition[5] by connecting the cognitive-denotative and affective-connotative meaning of concepts as measured with semantic differential rating scales. We attend to several preliminary considerations to set the stage for introducing the somatic transform, beginning with a discussion of levels of analysis and the importance of language for analysis at the psychological level, followed by the introduction of two principles (inextricability and complementarity) that summarize the relation between affect and cognition and that are rendered explicit by the somatic transform. With these background discussions in hand, we present affect control theory (ACT) and its Bayesian extension (BayesACT), which conjoin social psychological theory with computational tools to make the relation between affect and cognition explicit by connecting the cognitive-denotative and affective-connotative meaning of concepts through the somatic transform. Following the introduction of ACT and BayesACT, we identify examples of the inextricability and complementarity of affect and cognition in the social science and neuroscience literatures to set the somatic transform of BayesACT in a larger theoretical context. We conclude with a discussion of how our psychological model might be implemented in a realistic neural model.

Levels of Analysis

We are not proposing a Cartesian duality in which mind and body, or consciousness and matter, are distinct ontological realities (Walach, 2020).[6] Rather, we maintain that mental activity is a complex system that can be analyzed at the psychological level of the mind or at the neurological level of the brain where mental activity is physically implemented or realized, and both can be useful for gaining a better understanding of human intelligence and consciousness. The fact that the affect–cognition distinction is fuzzier at the level of the physical brain does not invalidate the distinction at the phenomenological level of the mind (Duncan & Barrett, 2007; see also Barrett & Satpute, 2013). We therefore accept as a phenomenological given that humans psychologically experience a distinction between hot and cold cognitions or, alternatively, between affective states mediated or largely unmediated by cognitive processing.

There is nothing unscientific in maintaining that while the mind is ontologically grounded or realized in the brain, it has emergent properties that cannot easily be reduced to neuronal activity. However, many neuroscientists schooled in reductionism would be uncomfortable with the idea of emergence. In the words of Gazzaniga,

> [N]euroscientists dislike this kind of thinking. They cling to the idea that understanding the elementary parts of the nervous system will explain how the brain does its magic to produce the psychological states we all enjoy. Letting go of this idea seems dangerous to most of them, who fear that some kind of ghost will be snuck into the brain . . . [in contrast] scientists working on other complex issues have no problem with the idea of emergence. Physicists, chemists, and biologists all know about it. (2010, pp. 1–2)

Among the most important of emergent properties is the reflective and experiential nature of consciousness, the elusive explanation of which has become known among neuroscientists as the “hard problem of consciousness” (Chalmers, 1995). How and why does our subjective or phenomenological experience arise out of the cognitive processing of auditory and visual information? Why and how do we have an inner life in which we can entertain images and thoughts, or experience emotions? And why and how do we experience ourselves as the locus of these experiences? Despite all its achievements, neuroscience has yet to provide a satisfactory and universally accepted answer.

Language

We live in a symbolic as well as a physical environment. While the medium of mental processing at the level of the physical brain is neuronal, the medium at the level of the psychological mind is largely symbolic. Without denying the importance of visual images, it is likely that mind and reflective consciousness emerged in the human species in step with the acquisition of language (Mead, 1934). Thus, approaching the study of mind and consciousness through the medium of language would appear to be the most appropriate and effective method for studying human subjectivity.[7]

Following Osgood et al. (Osgood, 1969; Osgood et al., 1957, 1975), we distinguish between the cognitive-denotative and affective-connotative meaning of language, between the labeling of objects and the affective associations evoked by labeling—“the ‘feeling tones’ of concepts as part of their total meaning” (Osgood, 1969, p. 195). Both kinds of meaning are essential to human communication. “Just as it is not possible to use words without communicating affective connotation, it is also not possible to understand language based on denotation alone” (Russell, 2003, p. 1199).

Compared to cognitive-denotative meaning, affective-connotative meaning is dimensionally simple, comprising feelings of evaluation (E), potency (P), and activity (A)—EPA. This Euclidean elegance provides a portal into the dimensionally complex cognitive-denotative world of meaning. “Osgood’s evaluation, potency, and activity [are] not simply . . . dimensions of spoken and written words . . . they are the hidden language, the affective Rosetta stone that allows mind and body to communicate” (Clore & Pappas, 2007, p. 338; see also Clore & Ortony, 2000). As discussed in what follows, ACT and

BayesACT employ semantic differential bipolar scales based on the EPA structure of affective-connotative meaning to measure cultural sentiments for social concepts.

EPA space describes a cognitively appraised affect handled by networks mainly in the prefrontal cortex, and this is not the same as "core affect" (Duncan & Barrett, 2007; Russell, 2003), which is handled by a different functional network in the brain and is much closer to interoceptive signals from the body (Beer, 2017). While Beer opines that few studies have "the aim of understanding the role of emotion in social cognition" (2017, p. 3), there is a large literature on the sociology of emotions that addresses precisely this issue.

The Inextricability and Complementarity of Affect and Cognition

On the basis of experimental evidence suggesting that subjects can make affective preferences among stimuli presented below the threshold of cognitive awareness, Zajonc famously concluded that "affect and cognition are separate and partially independent systems" (1984, p. 117), precipitating a heated exchange with Lazarus in the 1980s that became known at the time as the "primacy of cognition versus affect debate" (Zajonc, 1968, 1980, 1984; see also Lazarus, 1984; Zajonc, 2000).

Our position on the issue of primacy and independence of affect and cognition strikes a balance, specifying an intersection while allowing for partial independence. Moreover, we view the relation between cognition and affect as a reciprocal process, rendering the question of primacy or temporal-causal order a moot point. This same idea has been widely suggested in the literature (Forgas, 2008; Lazarus, 1984; Mook, 1987; Turner, 2009). The reciprocity of affective and cognitive processes is also a core assumption of affect control theory (ACT) and BayesACT discussed in the following lines.

This view of the relation between cognition and affect can be expressed as two principles (MacKinnon, 1994): (a) the principle of inextricability proposes that cognition and affect are overlapping constituents or simultaneous processes of the mind or consciousness rather than completely independent systems, a matter of relative preponderance where mind or consciousness at any given moment can be predominantly cognitive or predominantly affective or anywhere in between; and (b) the principle of complementarity proposes that, as only partially independent constituents or processes, both cognition and affect are necessary for an adequate understanding of the human mind. While the principle of inextricability is an ontological statement about the reality of the human mind as currently understood, the principle of complementarity is an epistemological implication of this ontological view.

At its most general level, the principle of complementarity proposes that what appear to be mutually exclusive phenomena are manifestations of some other underlying phenomenon, and that both expressions are necessary to understand it. The physicist Niels Bohr (1950) advanced the principle in the 1920s to explain the contradictory images evoked by the wave–particle duality of subatomic phenomena. The psychologist William James (1890) had developed the principle much earlier to reconcile, in his terms, the substantive and transitive parts of thought—the semantic denotations and syntactic connections of language with its inflections and affective overtones. And since both are essential components of human thought or consciousness, the principle of complementarity must be invoked in psychology just as it has been in quantum physics (Stephenson, 1986a, 1986b).

Because inextricability is a continuum rather than a dichotomy, one can distinguish between "harder" and "softer" views of the principle. Those adopting a harder view sometimes dismiss the distinction between cognition and affect altogether, even at the phenomenological level of the mind (Eder et al., 2007b). Those adopting a softer view, as we do, emphasize an overlap of cognitive and affective processing in the human mind while allowing that these processes may be partially distinct and independent systems (e.g., Storbeck & Clore, 2007).[8]

The inextricability and complementarity of affect and cognition are foundational assumptions of ACT and BayesACT. The cognitive-denotative and affective-connotative meanings of concepts are inextricable because one can be recovered from the other. At the same time, the two are complementary because they refer to the same deeper underlying reality and both are necessary to fully understand this deeper reality. And as expounded next, the somatic transform introduced in this article represents inextricability and complementarity explicitly in a simple energy potential connecting the cognitive-denotative and affective-connotative representations within the framework of Bayesian affect control theory.

Following an introduction to ACT and BayesACT, in the next two sections we identify examples of inextricability and complementarity in the social science and neuroscience literatures. This review locates the somatic transform of BayesACT in a larger theoretical context.

Affect Control Theory (ACT)

Affect control theory (ACT; Heise, 1979, 2007; MacKinnon, 1994, 2015; MacKinnon & Heise, 2010; Robinson & Smith-Lovin, 2018; Smith-Lovin & Heise, 1988) was developed from an integration of Osgood's semantic differential measurement model (Osgood, 1969; Osgood et al., 1957, 1975), Gollob's (1968) actor–behavior–object (ABO) model of impression-formation, and Powers's (1973) perceptual control model of behavior, blending these elements from psychology with insights from action theory and symbolic interactionism from sociology.

ACT measures the affective-connotative meaning of social concepts with three bipolar semantic differential scales corresponding to the EPA structure of affective-connotative meaning established by Osgood et al. (1957) The EVALUATION scale is anchored by "*bad, awful*" on one end to "*good, nice*" on the other end; the POTENCY scale, by "*small, weak, powerless*" to "*big, strong, powerful*"; and the ACTIVITY scale, by "*slow, old, quiet*" to "*fast, young, noisy*." Each scale ranges from infinitely "*bad, awful*" (−4.3) to infinitely "*good, nice*" (+4.3), with actual values generally falling between −3 and +3, and where a

+2 or −2 is considered a large (positive or negative) value. In ACT, individual scores on EPA scales are averaged to estimate the cultural sentiments of concepts.

Because EPA scales serve as generalized attitude scales (Osgood et al., 1957), they allow for the measurement of all kinds of social objects—social identities, interpersonal behaviors, social settings, social characteristics, personality traits, and emotions—on a single, common metric. Cultural surveys[9] of EPA semantic differential ratings of social objects yield a set of samples from a population distribution in sentiment space, which can be estimated parametrically (e.g., as the mean and variance of a normal distribution). Lists of mean EPA ratings of social objects mapping cognitive-denotative labels to cultural sentiments are referred to as "dictionaries" in ACT and have been constructed for Canada, China, Germany, Japan, Morocco, Northern Ireland, and the United States.[10]

Moreover, EPA scales provide a mathematically coherent metric, enabling the transformation of one type of phenomenon into another employing ACT models programmed in Interact,[11] a computer program for simulating social interactional events and their affective and cognitive outcomes. These models implement three fundamental principles of ACT: (a) affective reaction, (b) affect control, and (c) reconstruction.

1. The affective reaction principle proposes that people respond affectively to events, experiencing transient impressions or feelings for the actors, behaviors, and object-persons involved in them. Transient impressions are estimated by empirically derived impression-formation equations. The equation for actor evaluation, for example, reveals that good acts generate transient impressions of morality for the actor; powerful acts generate impressions of potency; and lively acts, impressions of liveliness. Being the recipient of a negatively evaluated action makes the object-person in an event decline in evaluation, and so on.
2. The affect control principle proposes that people construct and interpret events to minimize the discrepancy (called deflection in ACT) between internalized cultural sentiments and the transient feelings produced by events or, equivalently, to confirm cultural sentiments for the identities and actions involved in them. Thus, while the affective reaction principle pertains to impression-formation, the affect control principle applies to impression-management.
3. The reconstruction principle proposes that when people cannot minimize deflection through restorative actions, they attempt do so by reidentifying event participants by labeling them or by attributing explanatory traits and affective moods. Attributions are predicted from attribution equations, which are derived from empirical EPA ratings of identities modified by traits, moods, and other attributions.

Finally, ACT contains a theory of emotions, which proposes that emotions signal how an interaction is unfolding with respect to the identity-confirming or disconfirming impressions created for its participants. Empirically derived emotion equations in ACT reveal that emotions are a product of transient impressions and their deflection from fundamental cultural sentiments, thus serving as an "error signal" of sorts that can be communicated among participants to ensure smooth interaction through the selection and confirmation of situationally appropriate identities.

ACT unifies cognitive and affective processing by establishing a fundamental connection between the cognitive-denotative representations of the social environment and the affective-connotative representations of the sentiments or feelings with which they are associated. For example, when one perceives a person in a white coat in a hospital, a cognitive-denotative impression of this person may be formed and represented with the symbol "doctor," which has an associated cultural sentiment in three-dimensional EPA affective space. A doctor, for example, usually evokes feelings of goodness, strength, and modest activity, which is reflected in the EPA profile for this identity EPA profile (E: 2.7; P: 3.0; A: 0.23).[12] And given any affective-connotative (EPA) vector, a cognitive-denotative label can be assigned using a simple nearest neighbour method. For example, the closest label to the numerical EPA profile (−1.0; 2.0; 2.0) is "politician," with an EPA profile of (−0.9; 2.3; 1.5) at a Euclidean distance of 0.59 from this numerical EPA profile.

As stated before, EPA ratings yield a set of samples from a distribution in the sentiment space that represents the cultural group being measured. This distribution can be estimated parametrically (e.g., as the mean and variance of a normal distribution). However, ACT employs only the mean of EPA ratings to link cognitive-denotative and affective-connotative meaning, and only a single identity for each participant in analysis of actor–behavior–object events. This is restrictive because, in many situations, cognitive-denotative identities may not be clear, and sentiments may be more or less precisely defined. In our hospital example, a person in a white coat could be a nurse or a doctor (but probably not a patient). A person appraising this situation may be uncertain about the identity of the white coat wearer and might conclude that the identities of doctor or nurse are equally likely. Furthermore, while doctors are seen as more powerful than nurses, there may be many cultural experiences of powerful nurses and weak doctors, and cultural sentiment surveys may reveal this ambiguity. For example, the variance in the power ratings for a doctor is 1.4, while for nurse, with an EPA profile (2.9; 1.9; 0.7), the variance is significantly higher (2.4), indicating greater cultural uncertainty about the affective meaning of a nurse on the power dimension. A person observing the white coat wearer doing something powerful such as ordering someone to do something might shift her beliefs about the identity of this person, concluding perhaps that this person is much more likely to be a doctor than a nurse. The amount by which this belief is shifted will depend on the ambiguity of the cultural sentiments for the identities. For example, if doctors were more powerful in the mean, but with a very large variance, then an observation of a powerful behavior would decrease the shift, so the person might conclude that the white coat wearer is only a bit more likely to be a doctor. These deficits in ACT are addressed by BayesACT, to which we now turn.

Table 1. Mean and variance of sentiments for three identities.

Identity *x*	Mean M(x)			Variance y^2		
	E	P	A	E	P	A
Doctor	2.7	3.0	0.23	1.6	1.4	5.0
Nurse	2.9	1.9	0.7	1.9	2.4	3.6
Patient	0.64	−1.5	−1.3	1.2	1.1	2.0

Note. E = evaluation; P = potency; A = activity.

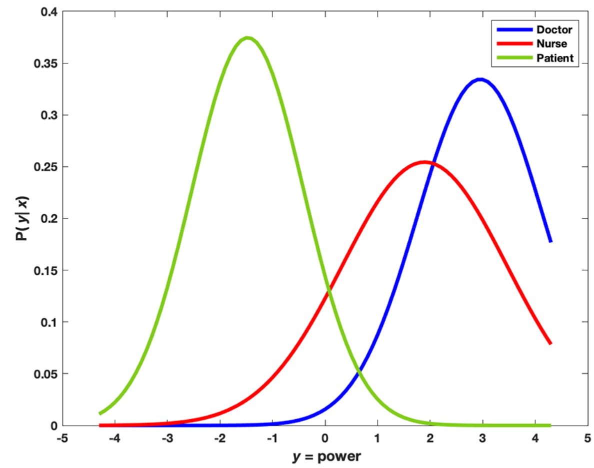

Figure 1. Plot of the normal connotative densities in y = power for three identities as estimated from the Georgia 2014 survey (Smith-Lovin et al., 2016).
Note. Compute code for generating this figure is available at bayesact.ca.

BayesACT and the Somatic Transform

BayesACT[13] (Hoey et al., 2016; Schröder et al., 2016) generalizes ACT by explicitly representing the distribution over a sentiment, taking the variance in sentiment measurements into account and releasing ACT from the constraints of a single identity and a mean sentiment. BayesACT allows for the explicit representation of the connection between denotative and connotative representations through a potential function we label "the somatic potential."[14] The original BayesACT model presented in Hoey et al. (2016) did not include this function. We introduce the somatic potential in this article and show how it can be used as a simple, expressive model of the principles of inextricability and complementarity applied to the relation between affect and cognition, naturally allowing for the modelling of multiple identities and of ambiguous sentiments.

The somatic potential is defined in terms of an energy[15] function that measures the incoherence (difference) between the denotative state (e.g., doctor) and the connotative state (a distribution in the affective EPA space). Denoting denotative states with x and connotative states with y, we write this energy function[16] as follows:

$$E(x,y) = (y - M(x))^2/\gamma^2 \tag{1}$$

where $M(x)$ gives the mean of the connotative distribution for a particular x, and where the parameter γ is the (square root of the) variance in sentiments in the population.

The energy function can be turned into a probability distribution, $G(x, y)$ called the somatic transform that can be used to rank the likelihood of different connotative and/or denotative events:

$$G(x,y) = c\,e^{-E(x,y)} = c\,e^{-(y-M(x))^2/\gamma^2}. \tag{2}$$

where c is a normalization constant.[17] Thus, the more coherent x is with y, the closer $M(x)$ is to y, the smaller the energy, and the higher the probability, with "closer," "smaller," and "higher" all characterized by the variance parameter γ. That is, if γ is larger, then the same amount of coherence is achieved when $M(x)$ is further from y.

The somatic transform can be used to translate between connotative and denotative states in either direction. For some x, $G(x, y)$ is a function over the connotative states representing the sentiments for denotative state x. For example, $G(x, y)$ for each of three identities in a hospital setting (nurse, doctor, and patient) is a normal distribution[18] given by summary statistics—mean $M(x)$, variance γ^2—for each identity, as shown in Table 1. Plotting these as probability density functions over the connotative space of power (P) only, we obtain the visualization shown in Figure 1. Figure 1 shows that the identity of a patient is significantly less powerful than either nurse or doctor, and that the distribution for nurse is significantly more spread out, indicating more ambiguity in the cultural sentiments for nurse than for either of the other two identities.

The somatic transform can also be used to map from connotative to denotative states. For ome set of denotative labels, a probability distribution over these can be formed for each y by normalizing $G(x, y)$ so it sums to 1.0 over the denotative labels. For example, plugging $y = 2.0$ into Equation 2 using the variances γ from Table 1, we find the unnormalized probability densities for each of the three hospital identities considered before, as shown in the second column of Table 2. A normalization (division by the sum of the three numbers) yields a probability distribution over these three identities, as shown in the third column of Table 2.

Therefore, suppose a person in a hospital was appraising an unknown individual based on sentiments or feelings about her/his power. Referring to the previous example, the person may have observed this individual ordering someone to do something. Starting from only a sentiment of power ($P = 2.0$), the likely identities can be ranked probabilistically by the somatic transform such that the identity of patient is all but ruled out, while those of nurse and doctor are roughly equivalent, with nurse slightly outranking doctor.

To investigate further, we can compute this probability distribution for every value of y, as shown in Figure 2. As the connotative power value in y moves from low to high, the most likely identity (top curve in Figure 2) changes from patient (for $y < 0.1$) to nurse (between $0.1 < y < 2.1$) to doctor ($y > 2.1$).

To recap our hospital example, a person may appraise a situation in terms of identities based on stereotypes and other learned meanings of cultural artifacts (e.g., gender, dress).

Table 2. Unnormalized somatic transformed values for power sentiment $y = 2.0$ for each of three identities and normalized probability distribution over those three identities.

Identity x	Unnormalized $G(x, y = 2.0)$	Normalized probability of x
Patient	0.0017	0.0035
Doctor	0.23	0.4775
Nurse	0.25	0.5190

These appraisals take the form of denotative representations (e.g., identity labels, "doctor" or "nurse"). Appraisals may also be in terms of sentiments (power, evaluation, or activity). The somatic transform allows for these elements to be combined in a probabilistic (Bayesian) way such that the resulting appraisals take both denotative and connotative meanings of identities and their variances into account. The somatic transform explicitly represents this inextricability and complementarity in a simple energy potential connecting the two kinds of representations within the framework of Bayesian affect control theory.

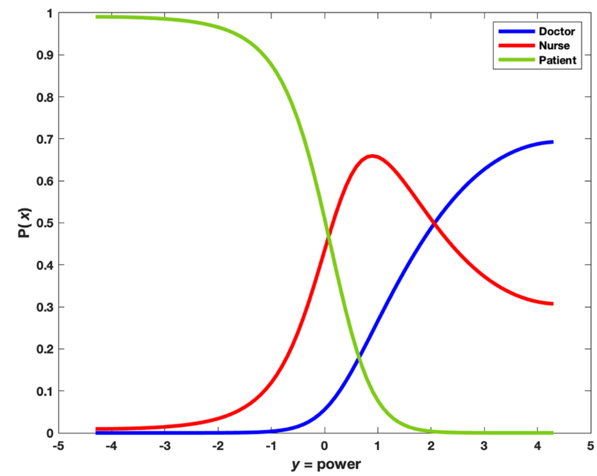

Figure 2. Probabilities over three identities as a function of the connotative value of power, y, as estimated through the somatic transform.

Note. Black dashed lines show the values of y at which the most likely identity changes. Compute code for generating this figure is available at bayesact.ca.

Inextricability and Complementarity in the Social Science and Neuroscience Literatures

As discussed before, the principles of inextricability and complementarity are foundational to ACT and BayesACT, and are rendered explicit by the somatic transform of BayesACT. In this section, we identify instances of these principles (explicit or implicit) in the social science and neuroscience literatures. This review sets the somatic transform in a more expansive theoretical context.

Social Science

Social scientists tend to adopt a softer position on inextricability, emphasizing the overlap of cognition and affect while allowing for a partial independence.

The inextricability of cognition and affect is embedded in the fundamental premise of cognitive appraisal theories of emotion, that cognitive processing is a necessary precondition for, or concomitant of, some important aspects of emotional experience. Although this theory has deep historical roots (Arnold, 1960; Lazarus, 1991; Mandler, 1975; Neisser, 1967), we restrict our discussion to the work of Clore, Ortony, and associates because of its affinity with ACT (Clore & Pappas, 2007).

In their earlier research on the "referential structure of the affective lexicon," Ortony et al. (1987, p. 351) concluded that "Mental conditions always have a significant cognitive component or a significant affective component, and sometimes both." This led them to differentiate between terms that have affect as their predominant referential focus (e.g., broken-hearted, thrilled) from other words that have both significant affective and cognitive components (e.g., confident, uninterested). And in a related study, Clore et al. (1987) found that people have difficulty in distinguishing pure affective states, such as angry and afraid, from intense cognitive ones, such as astonished and bewildered.

Inextricability is also a central theme in subsequent work by Clore, Ortony, and associates. In the case of specific emotions (e.g., angry, happy), Clore and Ortony (2000) assert that cognition is always implicated in emotional experience, because emotions always attach to objects, which are "necessarily represented, and representation is the essence of cognition" (2000, p. 53). In the case of undifferentiated affective states such as moods, Storbeck and Clore (2007) challenge the primacy, independence, and automaticity arguments for their alleged independence from cognition.

While emphasizing the effect of cognition on affect, Clore and associates also acknowledge the reciprocal effect of affect on cognition, "modulating and mediating" various cognitive processes (Storbeck & Clore, 2007, p. 1230). According to affect-as-information theory (Clore et al., 2018; Clore & Schnall, 2005; Schwarz & Clore, 1983), undifferentiated affective states such as moods play an important role in judgment and misattribution effects.

In addition to inextricability, the principle of complementarity is explicit in the writings of Clore, Ortony, and associates, as evidenced in their acknowledgement that "cognition and emotion are coming to be viewed as complementary rather than antagonistic processes" (Storbeck & Clore, 2007, p. 1225).

Like cognitive appraisal theory, psychological constructionist theory (Barrett, 2006, 2017a, 2017b; Barrett & Russell, 2015; Russell, 2003; Russell & Barrett, 1999) proposes that emotions emerge from a process of cognitive interpretation. While cognitive appraisal theory focuses on the interpretation of an external situation, however, psychological constructionist theory focuses on the interpretation of an internal sensory or affective state in making causal attributions to an external situation (Barrett, 2009; Gendron & Barrett, 2009).

The idea of inextricability is contained in the basic premise of the theory, according to which the specific emotions (happy, sad, etc.) are psychologically constructed from a combination of more general cognitive and affective primitives that are also ingredients of cognitions. The first primitive is "core affect"—that "neurophysiological state consciously accessible as the simplest raw (nonreflective) feelings evident in moods and emotions" (Russell, 2003, p. 148). Core affect is experienced as an ephemeral-to-enduring, free-floating affective event, an "affective substrate" (Duncan & Barrett, 2007, p. 1186) available to consciousness that is a precondition for first-person or phenomenological experience of the world.[19] Cognitive processes such as labeling and attribution conjoin with core affect in the psychological construction of discrete emotions.[20]

Psychological constructionist theory contains a harder view of inextricability for the neuronal level of the brain and a softer view for the psychological level of the mind. This is evident in the assertion that while the brain may not "respect" the distinction between cognition and affect, the distinction may be valid at the phenomenological level of the mind (Duncan & Barrett, 2007). The principle of complementarity is explicit in acknowledging that, at this level, "thinking and feeling . . . are actually two sides of the same coin" (Duncan & Barrett, 2007, p. 1202).

The ideas of inextricability and complementarity can also be found in sociological theories of emotion. For example, Franks (1989, 2006) adopts a harder version of inextricability for the neurological level of the brain and a softer version for the phenomenological level of the mind. While asserting that "Emotions are involved in many areas of the human brain and are tightly interwoven with structures of cognition, memory, and motivation" (2006, p. 51), Franks dismisses the affect–cognition divide at the level of the mind as "a fallacy of dualistic contrasts" (2006, p. 55). "Rather than thinking categorically," he argues, "it may be wiser to see emotion and cognition on a continuum with a very large middle ground" (2006, p. 60). And both inextricability and complementarity are contained in Franks's statement that a satisfactory resolution to the cognition–emotion issue "will depend on describing how they can be inextricably linked [inextricability] while capable of being in tension [complementarity]" (Franks, 2006, p. 55).

Neuroscience

Neuroscientists tend to adopt a harder version of inextricability, minimizing the distinction between cognition and affect, and emphasizing their overlap or functional integration in the brain. For example, Davidson (2003) includes in his list of "seven sins" to avoid in the study of emotion the assumptions that affect and cognition involve independent and separate neural circuitry, and that affect is mostly subcortical and cognition is mostly cortical. Commenting on the use of neuroimaging technology such as fMRI to locate cognitive and emotional processes in specific areas of the brain, Cacioppo et al. (2003) caution against equating "the beauty of a brain image" with "its psychological significance" (2003, p. 657), calling the idea of a close mapping of cognitive phenomena and underlying neural substrates a "category error" in cognitive neuroscience. And, more recently, Barrett and Satpute (2013) challenged the "faculty psychology" assumption that "emotional, social, and cognitive phenomena are realized in the operations of separate brain regions or brain networks" (2013, p. 1), proposing instead that findings from neuroimaging studies can be more effectively analyzed using a framework based on large-scale intrinsic networks—domain-general, functional information-processing networks—each of which comprises a number of nodes associated with more specific functions.

Other neuroscientists adopt a softer view of inextricability, acknowledging an intersection of cognition and affect without rejecting their distinction. This can be found in Damasio's theory of somatic markers (1994, 2003), which conjoins the bodily expression of emotion with the cognitive mechanisms of attention and working memory associated with the prefrontal cortices (for a recent update, see Poppa & Bechara, 2018). For Damasio, "emotion is played out under the control of both subcortical and neocortical structures" and "feelings are just as cognitive as any other perceptual image, and just as dependent on cerebral-cortex processing as any other image" (1994, p. 159).

A softer view of inextricability is explicit in LeDoux's often quoted assertion "that emotion and cognition are best thought of as *separate but interacting* [emphasis added] mental functions mediated by *separate but interacting* [emphasis added] brain systems" (1996, p. 69). The basis for distinguishing between cognitive and emotional states of consciousness, LeDoux argues,

> is not the system that represents the conscious content . . . but the systems that provide the inputs to the system of awareness. There is but one mechanism of consciousness and it can be occupied by mundane facts or highly charged emotions. (1996, p. 19)

This early proposition is echoed in the later development of a higher order theory of emotions (LeDoux & Brown, 2017; LeDoux & Hofmann, 2018), which proposes that "a general network of cognitions [GNC] underlies both cognitive and emotional states of consciousness" and that what distinguishes them is "the kind of inputs processed" (LeDoux & Hofmann, 2018, p. 69).[21]

Pessoa (2008, 2009) focuses on the interaction between emotion and cognitive processing made possible by the compact and extensive connectivity of neuronal networks in the brain. The idea of inextricability is explicit in his rejection of a distinction between cognition and affect based on cortical and subcortical brain areas. Because interaction logically implies some degree of distinction between cognition and affect, Pessoa's view of inextricability is of a qualified or softer nature. In later work, however, he suggests that we move beyond a focus on the interaction between cognition and emotion to understanding their integration in the brain, representing emotions in the brain as "functionally integrated systems that involve large-scale cortical-subcortical net-works that are sensitive to bodily signals" (Pessoa, 2018, p. 19; see also Pessoa, 2017).

Panksepp's view on the role of cognition in emotional experience differs significantly from that of other affective neuroscientists such as Damasio, LeDoux, and Pessoa. Based on extensive research with animal models, he argues forcibly for the evolutionary and functional primacy of affect in the mammalian (including human) brain, de-emphasizing the role of cognition in both unconscious and conscious affective experience. In doing so, however, he tries to walk a fine line between these two levels of consciousness. On the one hand, he holds that much of what transpires in the brain is automatic and unconscious, so that a substantial amount of emotional processing results from "neural networks that in themselves probably elaborate no conscious emotional feelings" (Panksepp, 2001, p. 142). On the other hand, he acknowledges the importance of conscious emotional feelings, but views them as "biological in their essential underlying form" (2001, p. 142), "emergent properties that are realized in the dynamic organizations of neuronal networks" (2001, p. 141). Thus, unlike other affective neuroscientists and cognitive psychologists, Panksepp maintains that neuronal processing in the subcortical brain is not only necessary but also sufficient to explain the conscious, subjective experience of emotional feelings, suggesting that we should explore the nature of conscious emotional feelings "subcortically where the cognitive light is dim but the affective light is bright" (2001, p. 159).

The cognitive neuroscience of emotion (hereafter CNE) proposed by Smith et al. (2018) contrasts with Panksepp's version of affective neuroscience. While Panksepp focuses on subcortical neural processes, CNE views emotional experience as a whole-brain phenomenon (Barrett, 2017a; Pessoa, 2017), a product of large-scale, intrinsic, domain-general, information-processing networks (see our previous discussion of Barrett & Satpute, 2013). The CNE approach of Smith et al. (2018) is embodied in their three-process model of sequential emotional experience: (a) affect response generation involving coordinated automatic changes in body state and cognitive attentional state; (b) representations of situational appraisals, emotion concepts, and body states; and (c) the consciousness of representations. Here, Smith et al. (2018) identify the large-scale, distributed neural system known as the global workspace network as responsible for selecting active representations for global broadcasting throughout the brain, allowing them to become consciously accessible for deliberation and decision-making. Thus, in contrast to Panksepp, the three-stage model of Smith et al. (2018) proposes that subcortical affective processes and cortical cognitive processes are inextricably involved in all three stages of emotional experience.[22]

Implementing the Somatic Transform Model as a Realistic Neural Model

We conclude this article with a suggestion of how the somatic transform might be implemented in a biologically realistic neural model, following the lead of the semantic pointer theory of emotion[23] proposed by Thagard and Schröder (2014; see also Kajic et al., 2019). This theory explains how the structural relations between the embodied features of emotion and the affective meaning of linguistic concepts represented by EPA (evaluation, potency, and activity) may be implemented in the biological neural networks of the brain, integrating the neurobiological and phenomenological levels of analysis in the study of emotion.

The connectionist networks of Thagard and Schröder (2014), among others, have the same motivation as the somatic transform: to build a formal model of the mutual influence of affect and cognition. In a connectionist network, this mutual influence is modeled as a weight vector in some population of neurons. In fact, there are biologically plausible ways of computing energy functionals of the weight vectors given a data set (e.g., by "Boltzmann machines"; Ackley et al., 1985), and minimizing an energy functional is the same as finding the posterior probability of the corresponding distribution over possible states. That is, if different neural networks representing cognition and affect are connected via distributions like the somatic transform that are characterized by a Boltzmann distribution, then it means they can be implemented in a neural network representation as in Thagard and Schröder (2014), or in a sample-based (Monte Carlo) representation as in BayesACT (Hoey et al., 2016). While the somatic transform and the conditional distributions could be implemented this way or in any of a number of other ways, such as the Emergent neural modelling system that implements some of the common types of connectionist networks (Aisa et al., 2008), the advantage of the sample-based approach used in Hoey et al. (2016) is a causal structure (a Bayesian network) that can be exploited in reasoning (finding the likely causes of events).

That is, BayesACT is a conceptual and mathematical model that is suitable for many different types of computational implementation, one of which is a Monte Carlo (stochastic simulation) method as in Hoey et al. (2016), while another uses a neural network that is biologically realistic. We are currently working on a third implementation as a nonbiologically realistic deep neural network. BayesACT and the somatic transform are theoretical models, but are computationally very difficult, requiring some form of approximation when implemented. This is a foundational problem in machine learning and artificial intelligence: the implementation of the model introduces a particular learning (model fitting) bias. Some of these implementations may be more useful in some contexts than in others. For example, one may be interested in examining how a real human brain trades off cognition and affect, and then the biologically realistic model would be most appropriate. Monte Carlo methods and deep neural networks, on the other hand, are not biologically realistic, may be more approximate but may also be much more computationally efficient, and would be more appropriate for artificial intelligence applications. Further, research on how to unify causal models such as Bayesian networks with deep neural networks is an ongoing effort (Rohekar et al., 2018).

Another advantage of the Monte Carlo methods from an artificial intelligence perspective is that they can make use of independencies encoded in the associated Bayesian network to accelerate the process of identifying the situation causally and computing strategies for action. On the other hand, it is not as clear how this causal aspect is implemented in a neural network, although there is some recent work on identifying important input features in a neural network, making them more interpretable, something considered essential in AI from an ethical perspective (Acien et al., 2019; Hooker et al., 2019; Stark & Hoey, 2021).

Summary and Conclusion

In summary, we have argued that cognition and affect are overlapping but partially independent systems of mind or consciousness. The principle of inextricability in its softer version represents this view, with the principle of complementarity as an epistemological implication. At the same time, we acknowledged that the distinction between cognition and affect becomes more problematic at the neurological level of the brain. The principle of inextricability in its harder form recognizes this reality. Although the focus of this article has been on the psychological level of analysis, we discussed these principles at the neurological level as well; and in the conclusion, we discussed how our psychological model of the somatic transform might be implemented as a realistic neural model in a variety of different ways for different contexts.

We approached the relation between cognition and affect at the symbolic level of language by capitalizing on the relation between the cognitive-denotative and affective-connotative meaning of concepts established by Osgood and associates. The somatic transform introduced in this article operationalizes this relation because it enables the mapping of cognitive-denotative and affective-connotative meaning onto one another, a mapping that is culturally shared. Connotative and denotative states are inextricable because one can be recovered from the other. At the same time, the two are complementary because they describe the same deeper underlying reality, and both are necessary to fully understand this deeper reality. Clearly, the connotative state will be of limited usefulness on its own; it needs to be translated into something concrete in the world, and, in particular, into concrete motor movements (behaviors). Perhaps less obvious, the denotative state by itself will also be of limited usefulness due to the computational difficulties it presents as environments grow more complex (FeldmanHall & Shenhav, 2019).

While important in itself, the somatic transform introduced in this article is part of a larger theory, BayesACT, a reformulation of affect control theory based on Bayesian probability theory and statistics. The affective dynamics of BayesACT take place at the representations of affective experience and conscious access levels in the three-process model of affective experience of Smith et al. (2018) discussed before, and would be involved in both subconscious and conscious cognitive appraisals. In this regard, Lane (2000) has formulated a five-level hierarchical model of emotional experience based on a continuum from unconscious to conscious processes. A Level 5 theory "yet to be formulated, would involve social cognition and would focus on how differentiated processes of self and other influences [sic] social behavior and the autonomic and neuroendocrine concomitants of emotional responses" (2000, p. 362). In view of the empirical basis of BayesACT in social cognitions and cultural sentiments, and its social psychological unit of analysis (interpersonal events), we believe that this theory approaches a Level 5 theory as defined by Lane (2000).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Neil J. MacKinnon https://orcid.org/0000-0002-3535-8606

Notes

1. These include Epstein's (1980, 1983, 1990) analytical-rational versus intuitive-experiential systems of information-processing, Haidt's (2001) rational versus intuitive processes in moral psychology, and Kahneman's (2011) systems of "slow" and "fast" thinking in behavioral (cognitive) economics. See Evans (2008) for a comprehensive review of dual-process models.
2. See Mandler (2002) for a comprehensive historical review of cognitive psychology, and Gendron and Barrett (2009) for a comparable review of the psychology of emotions.
3. For cognitive sociology, see Cerulo (2010, 2014), DiMaggio (1997), Lizardo (2014), and Vaisey (2009). For the sociology of emotions, see Bericat (2016), Smith-Lovin (1995), Stets and Turner (2014), Turner (1999, 2000, 2007, 2009), and Turner and Stets (2005, 2006).
4. In another article, we apply the somatic transform introduced for the first time here to the area of judgment and decision-making, employing examples of uncertainty and fairness, cognitive dissonance, and probability constraint models (Hoey et al., 2021).
5. Broadly defined, affect refers to psychological–physiological states such as emotions, feelings, sentiments, and moods; and cognition, to mental processes such as thought, attention, categorizing, knowledge acquisition, memory, reasoning, and so on. In this article, we funnel down to the affective-connotative and cognitive-denotative meaning of language as a way to access affect and cognition as broadly defined.
6. Walach (2020) proposes a "complementarist" ontology combining scientific materialism and phenomenological idealism, along with a corresponding dual complementarity epistemology of scientific empiricism and contemplative introspection.
7. This is exemplified by the importance of self-reports in social science and even in neuroscience, where they have been described as "the gold standard in studies of consciousness" (LeDoux & Hofmann, 2018, p. 67).
8. This distinction between softer and harder views of inextricability parallels LeDoux's (1996) distinction between "benign" and "not-so-benign" versions of cognitive science. In the benign version, the boundaries of cognition are expanded to include emotion on an equal level with thinking and reasoning. In the not-so-benign version, the mind is equated with cognition by "squeezing emotion into the traditional view of cognition . . . as thinking and reasoning" (1996, p. 68).
9. As opposed to population surveys, cultural surveys sample small numbers of "informants" from relatively homogeneous cultures and are designed to develop descriptive databases for the study of norms (see Heise, 2010).
10. These cultural data sets can be accessed at https://cs.uwaterloo.ca/~jhoey/research/ACTBackup/ACT/index.htm
11. Program Interact can be accessed at https://cs.uwaterloo.ca/~jhoey/research/ACTBackup/ACT/index.htm
12. EPA ratings in this article are from the Georgia 2015 Dictionary assembled by Smith-Lovin et al. (2016; see http://affectcontroltheory.org/). We will present EPA ratings as sets of three numbers assuming the order E, P, A.

13 The computer program for BayesACT can be accessed online at bayesact.ca. Compute code for generating Figures 1 and 2 is also available at bayesact.ca.

14 We use the term "somatic" here in the same sense of Damasio's somatic markers (1994, p. 173), as generally referring to visceral and nonvisceral attention guides providing a link between cognitive and affective states.

15 Energy here refers to statistical energy, which is a direct measure of a classifier's error on a data set, for example. Statistical energy is also related to probability, with more energetic systems being less probable as energy tends to dissipate in most physical systems according to the second law of thermodynamics.

16 Here, we work with a one-dimensional (potency) somatic potential and transform, for ease of exposition. In BayesACT, the transform is three-dimensional (EPA) in y, and so γ is a three-dimensional covariance matrix.

17 This form is known as the "Boltzmann distribution" and is only one common possibility for transforming energy (incoherence) into probability. We return to this term in the concluding section of this article.

18 The assumption of normality (that the population's sentiments are normally distributed) is used here, but may not be a good fit to the data for some cases. For example, sentiments about politically or socially controversial identities (e.g., abortionist) may have two modes in valence (E). In the Georgia 2014 data set (Smith-Lovin et al., 2016), this particular identity has a mean of $E = -1.1$, and a variance in E of 2.1, but is not well fit by a normal distribution, and would be more suitably modeled with a mixture of 2 or 3 Gaussians. In BayesACT, nonparametric representations are able to represent any degree of multimodality, given sufficient samples in a survey.

19 The second primitive in Russell's (2003) theoretical framework is the perception of the affective quality of external objects or stimuli that may affect a person's core affect, which by itself is a "cold" cognitive event "made hot by being combined with a change in core affect" (Russell, 2003, p. 148).

20 From this perspective, discrete emotions are not "natural kinds" but rather psychological constructions from the application of a person's conceptual knowledge about emotion to categorize a current state of core affect (Barrett, 2006; for an opposing view, see Adolphs, 2017).

21 In the case of emotional states such as fear, subcortical circuits provide lower order, nonconscious inputs that combine with cortical neural inputs in the GNC to bring about conscious emotional experience. Feedback from physiological and behavioral bodily response, emphasized by Damasio, also affects processing in the GNC and in arousal and survival circuits, creating feedback loops that maintain the reaction. The incorporation of a concept of self into their theory allows for self-awareness as a core and essential component of emotional experience (LeDoux & Brown, 2017).

22 In view of the preceding discussion, it would appear that Panksepp's theory of affective neuroscience and the CNE theory of emotion of Smith et al. (2018) are irreconcilable. However, advocates of each theory have collaborated on an attempt at rapprochement and a search for common ground (Panksepp et al., 2017).

23 Proposed by Eliasmith (2013), semantic pointers are interrelated patterns of neural activity generating complex conceptual hierarchies from sensorimotor representations to culturally constructed linguistic symbols.

References

Acien, A., Morales, A., Vera-Rodriguez, R., Bartolome, I., & Fierrez, J. (2019). Measuring the gender and ethnicity bias in deep models for face recognition. In R. Vera-Rodriguez, J. Fierrez & A. Morales (Eds.), *Lecture Notes in Computer Science: Vol. 11401. Progress in pattern recognition, image analysis, computer vision, and applications* (pp. 584–593). Springer. https://doi.org/10.1007/978-3-030-13469-3_68

Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for Boltzmann machines. *Cognitive Science*, *9*, 147–169.

Adolphs, R. (2017). How should neuroscience study emotions? By distinguishing emotion states, concepts, and experiences. *Social Cognitive and Affective Neuroscience*, *12*, 24–31.

Aisa, B., Mingus, B., & O'Reilly, R. (2008). The emergent neural modeling system. *Neural Networks*, *21*, 1146–1152.

Arnold, M. B. (1960). *Emotion and personality: Vol. 1. Psychological aspects*. Columbia University Press.

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

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

Barrett, L. F. (2017a). *How emotions are made*. Houghton Mifflin Harcourt.

Barrett, L. F. (2017b). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*, 1–23. https://doi.org/10.1093/scan/nsw154

Barrett, L. F., & Russell, J. A. (Eds.). (2015). *The psychological construction of emotion*. Guilford Press.

Barrett, L. F., & Satpute, A. (2013). Large-scale brain networks in affective and social neuroscience: Towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology*, *23*, 361–372. https://doi.org/10.1016/j.conb.2012.12.012

Beer, J. S. (2017). Current emotion research in social neuroscience: How does emotion influence social cognition? *Emotion Review*, *9*, 172–180. https://doi.org/10.1177/1754073916650492

Bericat, E. (2016). The sociology of emotions: Four decades of progress. *Current Sociology*, *64*, 491–513. https://doi.org/10.1177/0011392115588355

Bohr, N. (1950). On the notions of causality and complementarity. *Science*, *111*, 51–54.

Burke, P. J., & Reitzes, D. C. (1991). An identity theory approach to commitment. *Social Psychology Quarterly*, *54*, 239–251.

Burke, P. J., & Stets, J. E. (2009). *Identity theory*. Oxford University Press.

Cacioppo, J. T., Berntson, G. G., Lorig, T. S., Norris, C. J., Rickett, E., & Nusbaum, H. (2003). Just because you're imaging the brain doesn't mean you can stop using your head: A primer and set of first principles. *Journal of Personality and Social Psychology*, *85*, 650–661.

Cerulo, K. A. (2010). Mining the intersections of cognitive sociology and neuroscience. *Poetics*, *38*, 115–132.

Cerulo, K. A. (2014). Continuing the story: Maximizing the intersections of cognitive science and sociology. *Sociological Forum*, *29*, 1012–1019.

Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2*, 200–219.

Clore, G. L., & Ortony, A. (2000). Cognition in emotion: Always, sometimes, or never? In L. Nadel, R. Kane, & G. L. Ahern (Eds.), *The cognitive neuroscience of emotion* (pp. 24–61). Oxford University Press.

Clore, G. L., Ortony, A., & Foss, M. (1987). The psychological foundations of the affective lexicon. *Journal of Personality and Social Psychology*, *53*, 751–766.

Clore, G. L., & Pappas, J. (2007). The affective regulation of social interaction. *Social Psychology Quarterly*, *70*, 333–339.

Clore, G. L., Schiller, A. J., & Shaked, A. (2018). Affect and cognition: Three principles. *Current Opinion in Behavioral Sciences*, *19*, 78–82.

Clore, G. L., & Schnall, S. (2005). The influence of affect on attitude. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *Handbook of attitudes* (pp. 437–489). Lawrence Erlbaum.

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

Damasio, A. R. (2003). *Looking for Spinoza: Joy, sorrow, and the feeling brain*. Harcourt Brace.

Davidson, R. J. (2003). Seven sins in the study of emotion: Correctives from affective neuroscience. *Brain and Cognition*, *52*, 129–132.

DiMaggio, P. (1997). Culture and cognition. *Annual Review of Sociology*, *23*, 263–287.

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

Eder, A. B., Hommel, B., & De Houwer, J. (2007a). On the relation between cognition and affect [Special issue]. *Cognition and Emotion*, *21*. https://doi.org/10.1080/02699930701437386

Eder, A. B., Hommel, B., & De Houwer, J. (2007b). How distinctive is affective processing? On the implications of using cognitive paradigms to study affect and emotion. *Cognition and Emotion*, *21*, 1137–1154.

Eliasmith, C. (2013). *How to build a Brain: A neural architecture for biological cognition*. New York: Oxford University Press.

Epstein, S. (1980). The self-concept: A review and the proposal of an integrated theory of personality. In E. Staub (Ed.), *Personality: Basic issues and current research* (pp. 82–132). Prentice-Hall.

Epstein, S. (1983). The unconscious, the preconscious, and the self-concept. In J. Suls & A. Greenwald (Eds.), *Psychological perspectives on the self* (*Vol. 2*, pp. 219–247). Lawrence Erlbaum Associates.

Epstein, S. (1990). Cognitive-experiential self-theory. In L. A. Pervin (Ed.), *Handbook of personality theory and research* (pp. 165–192). Guilford.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278.

FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world. *Nature Human Behaviour*, *3*, 426–435.

Forgas, J. P. (2008). Affect and cognition. *Perspectives on Psychological Science*, *3*, 94–101.

Franks, D. D. (1989). Alternatives to Collins' use of emotion in the theory of ritualistic chains. *Symbolic Interaction*, *12*, 97–101.

Franks, D. D. (2006). The neuroscience of emotions. In J. E. Stets & J. H. Turner (Eds.), *Handbook of the sociology of emotions* (pp. 39–62). Springer.

Gazzaniga, M. (2010). Neuroscience and the correct level of explanation for understanding mind. *Trends in Cognitive Science*, *14*, 291–292. https://doi:10.1016/j.tics.2010.04.005

Gollob, H.F. (1968). Impression formation and word combination in sentences. *Journal of Personality and Social Psychology*, *10*, 341–353.

Gendron, M., & Barrett, L. F. (2009). Reconstructing the past: A century of ideas about emotion in psychology. *Emotion Review*, *1*, 316–339.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgement. *Psychological Review*, *108*, 814–834.

Heise, D. R. (1979). *Understanding events: Affect and the construction of social action*. Cambridge University Press.

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

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

Hoey, J., MacKinnon, N. J., & Schröder, T. (2021). *Denotative and connotative control of uncertainty: A computational dual-process model*. Judgment and Decision Making, 16, 505–550.

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. http://doi.org/10.1016/j.artint.2016.09.004

Hooker, S., Erhan, D., Kindermans, P. J., & Kim, B. (2019). A benchmark for interpretability methods in deep neural networks. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alche'-Buc, & E. Fox (Eds.), *Advances in neural information processing systems* (pp. 9737–9748). Neural Information Processing Systems Foundation.

James, W. (1890). *The principles of psychology*. Holt.

Johnson-Laird, P., & Oatley, K. (1992). Basic emotions, rationality, and folk theory. *Cognition and Emotion*, *6*, 201–223.

Kahneman, D. (2011). *Thinking, fast and slow*. Doubleday.

Kajic, I., Schroeder, T., Stewart, T. C., & Thagard, P. (2019). The semantic pointer theory of emotion: Integrating physiology, appraisal, and construction. *Cognitive Systems Research*, *58*, 36–53.

Lane, R. D. (2000). Neural correlates of conscious emotional experience. In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 345–370). Oxford University Press.

Lawler, E., Thye, S., & Yoon, J. (2009). *Social commitments in a depersonalized world*. Russell Sage.

Lawler, E., Thye, S., & Yoon, J. (Eds.). (2015). *Order on the edge of chaos: Social psychology and the problem of social order*. Cambridge University Press.

Lazarus, R. S. (1984). On the primacy of cognition. *American Psychologist*, *39*, 124–129.

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press.

LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life*. Simon & Schuster.

LeDoux, J. E., & Brown, R. (2017). A higher-order theory of emotional consciousness. *Proceedings of the National Academy of Sciences of the USA*, *114*, E2016–E2025. https://doi.org/10.1073/pnas.1619316114

LeDoux, J. E., & Hofmann, S. G. (2018). The subjective experience of emotion: A fearful view. *Current Opinion in Behavioral Sciences*, *19*, 67–72. https://doi.org/10.1016/j.cobeha.2017.09.011

Lizardo, O. (2014). Beyond the Comtean schema: The sociology of culture and cognition versus cognitive social science. *Sociological Forum*, *29*, 983–989.

MacKinnon, N. J. (1994). *Symbolic interactionism as affect control*. State University of New York Press.

MacKinnon, N. J. (2015). *Self-esteem and beyond*. Palgrave Macmillan.

MacKinnon, N. J., & Heise, D. R. (2010). *Self, identity, and social institutions*. Palgrave Macmillan.

Mandler, G. (1975). *Mind and emotion*. Wiley.

Mandler, G. (2002). Origins of the cognitive (r)evolution. *Journal of the History of the Behavioral Sciences*, *38*, 339–353.

Mather, M., & Fanselow, M. (2018). On emotion–cognition interactions [Special issue]. *Current Opinion in Behavioral Sciences*, *19*. https://doi.org/10.1016/j.cobeha.2017.09.011

Mead, G. H. (1934). *Mind, self, and society*. University of Chicago Press.

Mook, D. G. (1987). *Motivation: The organization of action*. Norton.

Neisser, U. (1967). *Cognitive psychology*. Prentice Hall.

Ortony, A., Clore, G. L., & Foss, M. (1987). The referential structure of the affective lexicon. *Cognitive Science*, *11*, 341–364.

Osgood, C. E. (1969). On the whys and wherefores of EPA. *Journal of Personality and Social Psychology*, *12*, 194–199.

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

Osgood, C. E., Suci, G. C., & Tannenbaum, P. H. (1957). *The measurement of meaning*. University of Illinois Press.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford University Press.

Panksepp, J. (2001). The neuro-evolutionary cusp between emotions and cognitions: Implications for understanding consciousness and the emergence of a unified mind science. *Evolution and Cognition*, *7*, 141–163.

Panksepp, J., Lane, R. D., Solms, M., & Smith, R. (2017). Reconciling cognitive and affective neuroscience perspectives on the brain basis of emotional experience. *Neuroscience and Biobehavioral Reviews*, *76*, 187–221.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*, 148–158.

Pessoa, L. (2009). Cognition and emotion. *Scholarpedia*, *4*, Article 4567. https://doi.org/10.4249/scholarpedia.4567

Pessoa, L. (2017). A network model of the emotional brain. *Trends in Cognitive Science*, *21*, 357–371.

Pessoa, L. (2018). Understanding emotion with brain networks. *Current Opinion in Behavioral Sciences*, *19*, 19–25.

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

Poppa, T., & Bechara, A. (2018). The somatic marker hypothesis: Revisiting the role of the "body-loop" in decision-making. *Current Opinion in Behavioral Sciences*, *19*, 61–66.

Powers, W.T. (1973). *Behavior: The Control of Perception*. Chicago: Aldine.
Robinson, D. T., & Smith-Lovin, L. (2018). Affect control theories of social interaction and self. In P. J. Burke (Ed.), *Contemporary social psychological theories* (2nd ed., pp. 139–165). Stanford University Press.
Rohekar, R. Y., Nisimov, S., Gurwicz, Y., Koren, G., & Novik, G. (2018). Constructing deep neural networks by Bayesian network structure learning. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems* (pp. 3047–3058). Curran Associates.
Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*, 145–172.
Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*, 805–819.
Schmidt, L. (2003). On affective neuroscience [Special issue]. *Brain and Cognition*, *52*. https://doi.org/10.1016/S0278-2626(03)00002-2
Schröder, T., Hoey, J., & Rogers, K. B. (2016). Modeling dynamic identities and uncertainty in social interactions: Bayesian affect control theory. *American Sociological Review*, *81*, 828–855.
Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.
Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review*, *74*, 29–39.
Smith, R., Kilgore, W. D. S., & Lane, R. D. (2018). The structure of emotional experience and its relation to trait emotional awareness: A theoretical review. *Emotion*, *18*, 670–692.
Smith-Lovin, L. (1995). The sociology of affect and emotion. In K. S. Cook, G. A. Fine, & J. S. House (Eds.), *Sociological perspectives on social psychology* (pp. 118–148). Allyn & Bacon.
Smith-Lovin, L., & Heise, D. R. (Eds.). (1988). *Analyzing social interaction: Advances in affect control theory*. Gordon & Breach.
Smith-Lovin, L., Robinson, D. T., Cannon, B. C., Clark, J. K., Freeland, R., Morgan, J. H., & Rogers, K. B. (2016). *Mean affective ratings of 929 identities, 814 behaviors, and 660 modifiers by University of Georgia and Duke University undergraduates and by community members in Durham, NC, in 2012–2014*. University of Georgia. http://affectcontroltheory.org/
Stark, L., & Hoey, J. (2021). The ethics of emotion in AI systems. In *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, March 2021 (pp. 782–793). New York (NY): Association for Computing Machinery (ACM).
Stephenson, W. (1986a). William James, Niels Bohr, and complementarity: I—Concepts. *The Psychological Record*, *36*, 519–527.
Stephenson, W. (1986b). William James, Niels Bohr, and complementarity: II—Pragmatics of a thought. *The Psychological Record*, *36*, 529–543.
Stets, J. E., & Turner, J. H. (Eds.). (2014). *Handbook of the sociology of emotions: Volume II*. Springer.
Storbeck, J., & Clore, G. L. (2007). On the interdependence of cognition and affect. *Cognition and Emotion*, *21*, 1212–1237.
Stryker, S. (1980). *Symbolic interactionism: A social structural version*. Benjamin/Cummings.
Swann, W. B., Griffen, J. J., Predmore, S. C., & Gaines, B. (1987). The cognitive–affective crossfire: When self-consistency confronts self-enhancement. *Journal of Personality and Social Psychology*, *52*, 881–889.
Thagard, P., & Schröder, T. (2014). Emotions as semantic pointers: Constructive neural mechanisms. In L. F. Barrett & J. A. Russell (Eds.), *The psychological construction of emotions* (pp. 144–167). Guilford.
Turner, J. H. (1999). Toward a general sociological theory of emotions. *Journal for the Theory of Social Behavior*, *29*, 132–162.
Turner, J. H. (2000). *On the origins of human emotions: A sociological inquiry into the evolution of human affect*. Stanford University Press.
Turner, J. H. (2007). *Human emotions: A sociological theory*. Routledge.
Turner, J. H. (2009). The sociology of emotions: Basic theoretical arguments. *Emotion Review*, *1*, 340–354.
Turner, J. H., & Stets, J. E. (2005). *The sociology of emotions*. Cambridge University Press.
Turner, J. H., & Stets, J. E. (Eds.). (2006). *Handbook of the sociology of emotions*. Springer.
Vaisey, S. (2009). Motivation and justification: A dual-process model of culture in action. *American Journal of Sociology*, *114*, 1675–1715.
Walach, H. (2020). Inner experience – Direct access to reality: A complementarist ontology and dual aspect monism support a broader epistemology. *Frontiers of Psychology*, *11*, Article 640.
Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*, 1–27.
Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*, 151–175.
Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist*, *39*, 117–123.
Zajonc, R. B. (2000). Feeling and thinking: Closing the debate over the independence of affect. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 31–58). Cambridge University Press.