Cogn Comput (2016) 8:1–14
DOI 10.1007/s12559-015-9374-4

Emotion: A Unified Mechanistic Interpretation from a Cognitive Architecture

Ron Sun[1] · **Nick Wilson**[1] · **Michael Lynch**[1]

Received: 26 October 2015 / Accepted: 1 December 2015 / Published online: 29 December 2015

Abstract This paper reviews a project that attempts to interpret emotion, a complex and multifaceted phenomenon, from a mechanistic point of view, facilitated by an existing comprehensive computational cognitive architecture—CLARION. This cognitive architecture consists of a number of subsystems: the action-centered, non-action-centered, motivational, and metacognitive subsystems. From this perspective, emotion is, first and foremost, motivationally based. It is also action-oriented. It involves many other identifiable cognitive functionalities within these subsystems. Based on these functionalities, we fit the pieces together mechanistically (computationally) within the CLARION framework and capture a variety of important aspects of emotion as documented in the literature.

Keywords Emotion · Cognitive architecture · Psychology · Computational

Issues of Emotion

The term "emotion" has come to denote a variety of somewhat different phenomena. It is also not entirely clear how one can identify something as an emotional as opposed to a non-emotional experience. Emotion is, to say the least, a complex and multifaceted phenomenon. Work in the fields of psychology and neuroscience has certainly contributed to a better understanding of aspects of emotion, but many fundamental issues, especially those concerning its mechanisms and processes, are yet to be understood. Computational models have also been developed, but they tend to be isolated models, not intrinsic to (or fully integrated into) an overall cognitive architecture that is psychologically tested and validated.

Human emotion manifests itself as a complex of experiential, behavioral, cognitive, psychological, and physiological characteristics, with many underlying mechanisms and processes. For instance, it has been variously characterized as physiological, cognitive, and/or goal-oriented (see, e.g., [7, 10, 18, 67]). It has also been argued (e.g., [51]) that emotion is the collective outcome of operations throughout a cognitive system. It should not be viewed as a unitary thing (although in engineering intelligent systems, a separate "emotion system" is often posited). That is, it is emergent, resulting from physiological reactions, action readiness, physical (external) actions, motivational processes, appraisal (with implicit and/or explicit processes), metacognitive processes, as well as decision making and reasoning of various forms (implicit or explicit).

Therefore, we have argued (see, e.g., [51, 61]) that human emotion should be computationally captured and explained by a comprehensive, generic, computational cognitive architecture [44, 49], with all its psychologically validated mechanisms and processes. That is, emotion should be captured and explained based on adequate representation of basic action decision making, reasoning, motivational, metacognitive, and other psychological processes, within a generic, comprehensive, computational model of the mind—a psychologically realistic computational cognitive architecture. These mechanisms and

This work has been supported in part by the ONR Grants N00014-08-1-0068 and N00014-13-1-0342 to the first author.

✉ Ron Sun
rsun@rpi.edu

[1] Cognitive Sciences Department, Rensselaer Polytechnic Institute, 110 Eighth Street, Carnegie 302A, Troy, NY 12180, USA

processes can capture the interaction of internally felt needs and external environmental factors in determining motivations and actions by individuals (hence emotion). They can capture the regularities within an individual in terms of behavioral manifestations at different times and with regard to different situations (social or physical), as well as behavioral variability [52].

Given the myriad of mechanisms and processes involved in emotion (e.g., physiological, cognitive, and goal-oriented processes), nothing short of a generic (and psychologically validated) computational cognitive architecture would be able to provide a comprehensive mechanistic interpretation of emotion. Even though a relatively small, somewhat standalone computational model may account for some aspects of emotion, it, by its very nature, will be limited in scope and thus less than completely satisfactory. Therefore, it becomes highly desirable that a model of emotion goes beyond relatively specialized computational models or those not fully integrated into an overall cognitive architecture.

Conversely, emotion, if it is a valid psychological construct, should be adequately accounted for by a psychologically realistic cognitive architecture, without any significant additions or modifications of mechanisms and processes within the cognitive architecture. A psychologically realistic cognitive architecture should, by nature, be comprehensive and include all necessary psychological mechanisms and processes. It should have all essential components of the mind (with their mechanisms and processes), such as various memory modules, various inference mechanisms, and various learning mechanisms, as its integral parts, and should also clearly identify their respective roles in the overall functioning. Otherwise, it would amount to a software tool or a programming language, which allows one to build whatever models that one wants to build but does not sufficiently specify the architectural elements of the mind. Thus, a psychologically realistic cognitive architecture should, in principle, have all the requisite mechanisms and processes in place to account for emotion. Ideally, a psychologically realistic computational cognitive architecture should be a model of emotion just by itself.

From this perspective, there are many open questions concerning emotion and the underlying psychological mechanisms and processes. How do these various processes involved in emotion take place? How do different mechanisms interact? For example, one might want to know:

- What role does motivation play in emotion (in detailed, mechanistic terms)?
- What role does emotion play in behavior (in detailed, mechanistic terms)?
- What roles do explicit and implicit processes play in emotion, respectively?
- How do implicit and explicit processes interact in emotional processing?
- How can emotion be regulated through metacognitive means?

With a detailed, psychologically realistic computational cognitive architecture, we have started to address these questions in a more tangible and better constrained way, utilizing the mechanisms and processes specified within the cognitive architecture, which may lead to unique, broad interpretations of emotion. An overview of this project is provided in the present article.

Below, first, a comprehensive framework in the form of a psychologically realistic cognitive architecture is outlined. Then, on that basis, a number of key issues concerning emotion are discussed. Based on the discussions, a model of emotion is outlined, and some details are sketched within the cognitive architecture.

A Comprehensive Framework Capable of Addressing Emotion

Overview of CLARION

CLARION is a cognitive architecture—a (relatively) comprehensive model of essential psychological mechanisms and processes, specified computationally. It has been described, justified, and psychologically validated extensively on the basis of psychological data in Sun [44, 45, 49]; see also Sun et al. [54, 55] and Helie and Sun [12]. For the sake of subsequent discussions, a quick sketch is provided below.

CLARION consists of a number of subsystems (see [49] for detailed arguments regarding the division of subsystems): the action-centered subsystem (the ACS), the non-action-centered subsystem (the NACS), the motivational subsystem (the MS), and the metacognitive subsystem (the MCS). The role of the action-centered subsystem is to control actions (regardless of whether the actions are for external physical movements or for internal mental operations), utilizing procedural knowledge. The role of the non-action-centered subsystem is to utilize declarative knowledge for information and inferences of various kinds. The role of the motivational subsystem is to provide underlying motivations for perception, action, and cognition (in terms of providing impetus and feedback). The role of the metacognitive subsystem is to monitor and regulate the operations of the other subsystems dynamically [49].

Each of these interacting subsystems consists of two “levels” of representations (i.e., a dual-representational structure), as theoretically posited in Sun [44]. Generally speaking, in each subsystem, the "top level" encodes *explicit* knowledge with associated explicit processes (using symbolic/localist representations), while the "bottom level" encodes *implicit* knowledge with associated implicit processes (using connectionist distributed representations; [36]). The two levels interact, for example, by cooperating in action decision making through integration of the action recommendations from the two levels of the ACS, respectively, as well as by cooperating in learning through bottom-up and top-down learning processes (in the ACS or in the NACS; [54, 55]); see Fig. 1 for a sketch of CLARION.

As has been pointed out before, existing theories tend to confuse implicit and explicit processes, hence the "perplexing complexity" [42]. In contrast, CLARION generally separates implicit and explicit processes in each of its subsystems. With this framework, CLARION can provide explanations of empirical findings in a wide range of domains (see, e.g., [12, 54, 55, 63]).

Another important characteristic of CLARION is that it embodies the belief that cognition is activity-based, action-oriented, and embedded in the world [44, 50]. Therefore, for example, the principle regarding reasoning in CLARION is: action first and reasoning in the service of action.

Yet another important characteristic of CLARION is its focus on the cognition–motivation–environment interaction [46, 49], as opposed to dealing only with cognition in its narrow sense.

Below, we will examine each of its subsystems in more detail, which will illustrate some of these points (see [44, 45, 49] for further details).

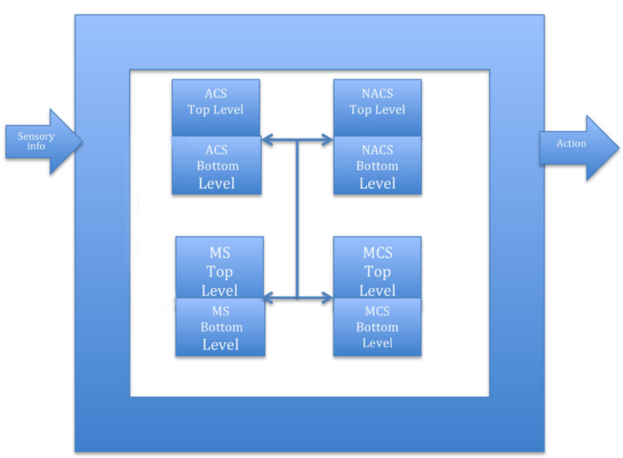

Fig. 1 Subsystems of the CLARION cognitive architecture. The major information flows are shown with arrows. ACS stands for the action-centered subsystem. NACS stands for the non-action-centered subsystem. MS stands for the motivational subsystem. MCS stands for the metacognitive subsystem. See the text for more explanations

Action-Centered Subsystem

The action-centered subsystem (the ACS) captures the action decision making of an individual when interacting with the world, involving procedural knowledge [44, 48].

In the ACS, the process for action selection is essentially the following: Observing the current (observable) state of the world, the two levels within the ACS (implicit or explicit) make their action decisions in accordance with their respective (procedural) knowledge, and their outcomes are "integrated." Thus, a final selection of an action is made and the action is then performed. The action changes the world in some way. Comparing the changed state of the world with the previous state somehow, the person learns. The cycle then repeats itself.

Thus, the overall algorithm for action decision making may be described as follows:

1. Observe the current input state x (including the current goal).
2. Compute in the bottom level the "value" of each of the possible actions (a_i's) associated with the current input state x: $Q(x, a_1)$, $Q(x, a_2)$,……,$Q(x, a_n)$. Stochastically choose one action according to these values.
3. Find out all the possible actions at the top level (b_1, b_2, …, b_m), based on the current input state x (which goes up from the bottom level) and the existing rules in place at the top level. Stochastically choose one action.
4. Choose an action, by stochastically selecting the outcome of either the top level or the bottom level.
5. Perform the action and observe the next input state y and (possibly) the reinforcement r.
6. Update implicit knowledge at the bottom level in accordance with an appropriate learning algorithm (e.g., Q-learning; more later), based on the feedback information.
7. Update explicit knowledge at the top level using an appropriate learning algorithm (e.g., the RER algorithm; more later).
8. Go back to Step 2.

In this subsystem, the bottom (implicit) level is implemented using neural networks involving distributed representations [36], and the top level is implemented using symbolic/localist representations.

For the bottom level, the input state (x) consists of the sensory input (environmental or internal), the current goal, and the working memory. All that information is important in deciding on an action. The input state is represented as a set of microfeatures. The output of the bottom level is the action choice, also represented as a set of microfeatures.

At the top level, "chunk" nodes are used for denoting concepts. A chunk node connects to its corresponding microfeatures at the bottom level (represented by a set of separate nodes constituting a distributed representation in the bottom level). At the top level, action rules connect chunk nodes representing conditions to chunk nodes representing actions. If the condition of an action rule is met, then the corresponding action is recommended.

At the bottom level of the ACS, with neural networks encoding implicit knowledge, actions are selected based on their Q values, which are the outputs of the neural networks. A Q value is an evaluation of the "quality" of an action in a given input state: $Q(x, a)$ indicates how desirable action a is in state x. At each step, given input state x, the Q values of all the actions (i.e., $Q(x, a)$ for all a's) are computed in parallel. Then the Q values are used to decide stochastically on an action to be performed, through a Boltzmann distribution of Q values (i.e., a softmax function):

$$p(a|x) = e^{Q(x,a)/\tau} / \sum_i e^{Q(x,a_i)/\tau}$$

where τ (temperature) controls the degree of randomness of action decision making and i ranges over all possible actions. (This is known as Luce's choice axiom; see [59].)

For learning implicit knowledge at the bottom level (i.e., the Q values), the Q-learning algorithm [59], which is a reinforcement learning algorithm, may be used. Q values are gradually tuned through successive updating, which enables reactive sequential behavior to emerge through trial-and-error interaction with the world [54, 59]. As a result of such learning, the Q values come to represent, roughly, the maximum cumulative reinforcement that can be received from the current point on, where reinforcement represents the fulfillment of needs and achievement of goals (as decided by the MS and the MCS; more later).

For learning explicit action rules at the top level with a bottom-up learning process, the *rule–extraction–refinement* (RER) algorithm utilizes information from the bottom level in learning rules at the top level; see Sun et al. [54] and Sun [44] for details. On the other hand, top-down learning goes in the opposite direction [44].

For stochastic selection of the outcomes of the two levels, at each step, with probability P_{BL}, the outcome of the bottom level is used. Likewise, with probability P_{RER}, the outcome from the RER rule set is used. Other components, if they exist, may also be included in the stochastic selection. There exists some psychological evidence for such intermittent use of rules [54].

Non-Action-Centered Subsystem

The non-action-centered subsystem (the NACS) is for dealing with declarative, or non-action-centered, knowledge [48, 49]. It stores such knowledge in a dual-representational form (the same as in the ACS): that is, in the form of explicit "associative rules" (at the top, level), and in the form of implicit "associative memory" (at the bottom level). Its operation is under the control of the ACS and it is in the service of action decision making usually.

At the bottom level of the NACS, associative memory networks encode implicit declarative knowledge. Associations are formed by mapping an input to an output (e.g., [36]).

At the top level of the NACS, explicit declarative knowledge is stored. As in the ACS, chunk nodes (denoting concepts) at the top level are linked to microfeatures represented at the bottom level. Additionally, at the top level, links between chunk nodes encode explicit associative rules (which may be learned in a variety of ways; [45, 49]).

As in the ACS, top-down or bottom-up learning may take place in the NACS, either to extract explicit knowledge at the top level from the implicit knowledge in the bottom level, or to assimilate the explicit knowledge of the top level into the implicit knowledge at the bottom level.

With the interaction of the two levels, the NACS carries out rule-based, similarity-based, and constraint-satisfaction-based reasoning. The overall operation of the NACS is as follows:

1. A directive is received by the NACS to initiate reasoning on a specified input.
2. Bottom-up and top-down activation propagate input activations to both levels of the NACS.
3. Associative reasoning is performed simultaneously at both levels:
   - 3.a. Associative memory networks propagate activations at the bottom level.
   - 3.b. Associative rules activate chunks at the top level.
4. Activations of the two levels are integrated at the top level.
5. At a set time limit or when no further conclusions can be inferred, the NACS returns chunks that were inferred. Otherwise, the process is reiterated (e.g., using the results of the previous iteration as inputs).

Details of the NACS will not be covered here because they are not needed in this paper (details can be found in, e.g., [12] or [49]).

Motivational Subsystem

The motivational subsystem (the MS) is concerned with why an individual does what he/she does [46]. The relevance of the MS to the ACS lies primarily in the fact that it provides the context in which goals and reinforcements for

Table 1 Primary drives, including approach-oriented and avoidance-oriented drives

Approach drives	Avoidance drives	Both
Food	Sleep	Affiliation and belongingness
Water	Avoiding danger	Similance
Reproduction	Avoiding unpleasant stimuli	Deference
Nurturance	Honor	Autonomy
Curiosity	Conservation	Fairness
Dominance and power		
Recognition and achievement		

the ACS are determined. It therefore influences the working of the ACS (and by extension, the working of the NACS).

A dual motivational representation is in place in the MS. The explicit *goals* (such as "*find food*"), which are essential to the working of the ACS as explained before, may be generated based on implicit *drives* (e.g., "*hunger for food*"). The explicit goals derive from, and hinge upon, implicit drives. For psychological justifications, see Sun [46, 49]; in this regard, see also Tolman [57], Maslow [25], and Wright and Sloman [66].

Among drives, *primary drives* are essential, most likely built-in (hard-wired) to a significant extent to begin with. Some sample *low-level* primary drives include *food, water,* and *reproduction*. Beyond those concerning mostly physiological needs, there are also *high-level* primary drives, for example, *dominance, fairness*, and *deference*, [30, 35, 46]. The primary drives (both low level and high level) are listed in Table 1. This set of primary drives has been extensively justified [46].[1]

Note that some of these primary drives are approach-oriented, while others are avoidance-oriented. This distinction has been argued by many before (e.g., [4, 42]). The approach system is sensitive to cues signaling rewards and results in active approach. The avoidance system is sensitive to cues of punishment and results in avoidance, characterized by anxiety or fear. The division between approach-oriented and avoidance-oriented drives provides an underlying structure for the approach and avoidance systems.

The processing of the drives involves a number of modules. In particular, the core module determines drive strengths (using a neural network) based on:

$$ds_d = gain_u * gain_s * gain_d * stimulus_d * deficit_d + baseline_d$$

where ds_d is the strength of drive d, $gain_d$ is the individual gain for drive d, $gain_u$ is the universal gain affecting all drives, $gain_s$ is the gain affecting all the drives of one type (e.g., the approach or the avoidance type), $stimulus_d$ is a value representing how pertinent the current situation is to drive d, $deficit_d$ indicates the perceived deficit in relation to drive d (which represents an individual's intrinsic sensitivity and inclination toward activating drive d), and $baseline_d$ is the baseline strength of drive d. The justifications for this mapping may be found in a variety of studies (e.g., [46, 56]). In particular, the multiplicative combination of $stimulus_d$ and $deficit_d$ has been argued for [46, 58].

Metacognitive Subsystem

The existence of the large number of drives and the goals resulting from them leads to the need for metacognitive control and regulation [9, 33]. In CLARION, the metacognitive subsystem (the MCS) is closely tied to the MS. The MCS monitors, controls, and regulates cognitive processes. Control and regulation may be in the forms of setting goals (which are then used by the ACS) on the basis of drives, generating reinforcement signals for the ACS for learning (on the basis of drives and goals), interrupting and changing ongoing processes in the ACS and the NACS, setting essential parameters of the ACS and the NACS, and so on [49].

Structurally, this subsystem may be divided into a number of functional modules, including among others:

- the goal module,
- the reinforcement module,
- the processing mode module,
- the input filtering module,
- the output filtering module,
- the parameter setting module.

Let us look into one module as an example. The goal module, in order to select a new goal, first determines goal strengths for all possible goals, based on information from the MS (e.g., the drive strengths) and the current sensory input. In the simplest case, the following calculation is performed:

$$gs_g = \sum_{d=1}^{n} relevance_{d,s \rightarrow g} * ds_d$$

[1] Briefly, this set of hypothesized primary drives bears close relationships to Murray's needs [30] and Reiss's motives [35]. The prior justifications of these frameworks may be applied, to a significant extent, to this set of drives as well [25, 30, 35, 46].

where gs_g is the strength of goal g, $relevance_{d,s \to g}$ is a measure of how relevant drive d is to goal g with regard to the current situation s (which represents the support that drive d provides to goal g), and ds_d is the strength of drive d as generated by the MS. Once calculated, the goal strengths are turned into a Boltzmann distribution and the new goal is chosen stochastically from that distribution. Arguments in support of goal setting on the basis of implicit motives (i.e., drives) may be found in, for example, Tolman [57], Wright and Sloman [66], and Sun [46].

Simulations Using CLARION

CLARION has been successfully applied to simulating, accounting for, and explaining a wide variety of psychological data. For example, a number of well-known skill learning tasks have been simulated and explained using CLARION that span the spectrum ranging from simple reactive skills to complex cognitive skills. Among them, some are typical implicit learning tasks (mainly involving implicit processes), while some others are high-level cognitive skill acquisition tasks (with significant presence of explicit processes). Simulations have also been done with reasoning tasks, metacognitive tasks, and motivational tasks. While accounting for various psychological data, CLARION provides explanations that shed new light on underlying processes (see, e.g., [12, 52, 54, 55]).

CLARION similarly provides explanations of, and clarifications to, issues related to emotion. To explore this aspect, we first need to address a number of foundational issues concerning emotion and then we will fit all the necessary pieces together to form a model of emotion.

Addressing Issues of Emotion within the Framework

Below, on the basis of the general CLARION framework briefly sketched above, we examine a number of issues fundamental to understanding emotion, before putting together a complete model. But two points should be emphasized first.

First, as mentioned earlier, there are reasons to believe that, psychologically, emotion is the collective outcome of operations throughout a cognitive system. It is emergent, resulting from mechanisms and processes associated with physiological reactions, action readiness, physical (external) actions, motivational processes, appraisal processes, metacognitive processes, as well as decision making and reasoning of various forms. Human emotion is the sum total of all of the above in particular circumstances [6, 16, 51]. Thus, in CLARION, emotion involves, for example, the ACS for actions, the NACS for reasoning, the MS for motivation, and the MCS for metacognitive regulation.

Second, a number of basic emotions have been identified by various researchers (e.g., [7]), for dealing with "fundamental life tasks." Different researchers may have identified somewhat different sets of basic emotions, but they usually include commonly discussed emotions such as anger and elation. These basic emotions need to be accounted for in particular in a mechanistic interpretation of emotion.

Emotion and Motivation

One natural hypothesis based on the CLARION framework, which is motivationally based, is that emotion is rooted in basic human motives or needs (i.e., drives) and their possible fulfillment [51]. In this regard, some other researchers, for example, Smillie et al. [42], Carver and Scheier [3], and Ortony et al. [31], also stressed the importance of motivation and expectation in emotion. Relatedly, Reisenzein [34] viewed emotions as propositional objects directed at states of affairs.

Thus, within the CLARION framework, emotions should be analyzed in terms of their motivational underpinnings. For example, it may be hypothesized within the CLARION framework that the emotion of *elation* is related to positive reward (including unexpected positive reward) and also, to a lesser extent, "expectation" of positive reward [13, 26, 51]. The intensity of elation may be (in part) related to the strengths of approach-oriented drives within the MS of CLARION (as discussed in section "A Comprehensive Framework Capable of Addressing Emotion").

For another example, the emotion of *anxiety* can be related to "expectation" of negative reward. The intensity of anxiety may be (in part) a function of the strengths of avoidance-oriented drives within the MS (see section "A Comprehensive Framework Capable of Addressing Emotion"). Smillie et al. [42] specifically identified the link between the avoidance system and anxiety. Carver and Sheier [3] and a number of others also made related points.

Furthermore, the emotion of *fear* may be due to "expectation" of more intense negative reward. The intensity of fear can be determined in a similar way as anxiety—(in part) as a function of avoidance-oriented drive strengths within the MS. In clinical psychology and psychophysiology research, generally speaking, there has been a lack of clear distinction between anxiety and fear [42]. Some demonstrated that what was constructed to represent "pure fear" situations was also a predictor of trait anxiety.

For yet another example, the emotion *anger* can be attributed to a mismatch between the "expectation" of a behavior from others and the actual behavior [26] when the actual behavior leads to more negative reward or less positive reward compared with the expectation. Computationally, the difference leads to high drive strengths within the MS for some avoidance-oriented and approach-oriented drives (e.g., the *fairness* drive).

Similar descriptions, on the basis of the CLARION framework, can be applied to other basic emotions as have been identified by various researchers (e.g., [7]). However, there may be some differences between colloquial usages of emotion terms (and the folk psychology behind them) and the interpretation here. Reinterpretation and clarification are necessary.

Emotion and the Implicit–Explicit Distinction

We need to examine the relationship between the implicit–explicit distinction and emotion. Psychological and neuroscience research suggested that emotional processes represented a more primitive mechanism (e.g., [20]). Experimental work indicated that emotional processing tended to quickly identify stimuli that were highly dangerous or beneficial (e.g., to an individual's survival). Emotions were often associated with hardwired and specific responses [7, 67]. Scherer (e.g., [39]) claims that emotion is complex, has multiple components, and often cannot be described by verbal labels. Lewis [22] claims that emotion emerges from a dynamic system.

Psychological and neuroscience research on implicit memory and implicit learning indicated the existence of distinct systems with distinct characteristics [8, 32], as embodied by CLARION (discussed earlier), some of which were analogous to the characteristics of emotion-processing mechanisms identified above. In particular, it was commonly believed that the emotion system was often faster and less differentiated, while the other, non-emotion system was slower and more deliberative [5, 67]. This distinction was similar to what was described as the distinction between explicit and implicit processes in general by, for example, Schacter [37], Reber [32], Sun [44], and Evans and Frankish [8]. Thus, this distinction with regard to emotion is consistent with the CLARION framework.

However, the separation of the emotional and the non-emotional processes is certainly limited. For instance, researchers have described a variety of appraisal processes, which rely, to some extent at least, on explicit processes [7, 10]. Appraisal is important in that it is involved in inducing an emotional state in reaction to a particular state of the world as perceived by an individual [10, 43]. This situation is somewhat analogous to similarly complex interaction between explicit and implicit processes, as studied, for example, by Sun et al. [55]; see also Sun [44] in this regard.

Such separation and interaction, with regard to emotion or with regard to implicit–explicit processes in general, are consistent with the CLARION framework as reviewed earlier. Therefore, the mechanisms and processes specified in CLARION, at both the top and the bottom level, may help to understand emotional processing. Emotional processing is mostly implicit (i.e., at the bottom level of CLARION), although not all implicit processes are emotional [5, 20, 67]. In CLARION, emotional processes involve various subsystems, as alluded to earlier. Emotional processing mainly occurs in the bottom levels of these subsystems [51]. However, explicit processes (at the top level) have a role too, for example, through performing cognitive appraisal (e.g., within the NACS; [10, 43]), or through affecting implicit processes in other ways [54, 55]. However, they are not the main locus of the experience of emotion according to CLARION.

Effects of Emotion

It has been observed that emotion involves (and has various effects on) perception, action, and cognition. We need to examine their involvement in emotion.

First, emotion is closely tied to action. On the basis of motivation, emotion leads to action usually. In fact, emotion manifests, to a significant extent, through actions. This means that, within CLARION, on the basis of the MS, emotion usually leads to actions by the ACS, involving both implicit and explicit processes of the ACS, with implicit processes being more fundamental (as discussed earlier [51]).

Furthermore, beyond being just related to action, emotion is fundamentally action-oriented. For example, Frijda [10], among others, indicated the importance of "action readiness" in emotional experience. As pointed out before, emotion involves physiological reactions, action readiness, physical (external) actions, motivational processes, appraisal processes, metacognitive processes, as well as decision making and reasoning of various forms. But, fundamentally, it involves first physiological reactions for the sake of action readiness and consequently physical (external) actions when circumstances permit them. Such actions, among other types, include external emotional expressions (such as facial expressions; [7]). Zajonc [67] and others argued for the physiological nature of emotion.

Second, emotion has various effects on perception. This phenomenon has been observed experimentally. For example, when in a state of anxiety, attention is heightened with regard to threatening stimuli. When in a state of positive affect, stimuli are more elaborately processed (e.g., [2, 27]).

Research has also shown the effects that emotion has on cognition [19, 39, 41]. Emotion may help to make the mind more adaptive. With emotion in place, one has at one's disposal both simple reflexive (i.e., fast emotional) responses (e.g., action readiness) and complex, elaborate cognitive processing, as well as their combination and interaction. Research has suggested that emotion involves and affects all functionalities of cognition, including attention, memory, learning, reasoning, and decision making. This means that, within CLARION, emotion involves and affects the ACS, the NACS, and the MCS, in addition to involving the MS as discussed earlier. Emotion involves both implicit and explicit processes in these subsystems, with implicit processes being more fundamental (as discussed earlier [51]).

The question is through exactly what mechanisms and processes emotion involves and affects these different functionalities such as attention, memory, learning, and reasoning. To address this question, a variety of computational models were proposed in the past, ranging from earlier ones such as Wright and Sloman [66] to more recent ones such as Marsella and Gratch [24]. However, most of these computational models are not psychologically validated through detailed comparisons with quantitative human data. Many of them espoused rather explicit processing (but with some exceptions); as such, they dealt with mostly a limited kind of emotion, which was not necessarily the most fundamental kind (as discussed before). Many of these models were also often standalone models (to a large extent at least, but with some exceptions), and as such, they were not fully integrated into the overall cognitive architecture in a psychologically realistic way. Thus, it appears that CLARION can potentially provide a more comprehensive, more integrative, more unified, and more psychologically realistic account through its generic mechanisms and processes resulting from, and compatible with, modeling various other cognitive–psychological functionalities (as will be examined later [51]).

Emotion Generation and Regulation

Emotion generation is accomplished through motivation, appraisal, and action [61]. Among these processes, motivation and action were addressed earlier, so we now look into appraisal.

For emotion generation, besides motivation and action, appraisal is important. A principle tenet of appraisal theory was that emotion was a result of "cognitive appraisal" (e.g., [10]). The model of Marsella and Gratch [24], for instance, implemented a form of appraisal theory. It suggested that, to adequately capture emotion, appraisal processes needed to rely on declarative knowledge and reasoning. Another model by Reisenzein [34], however, assumed that emotion arose when discrepancies were detected by continuously running, rapid, and automatic appraisal processes.

Within CLARION, both types of "appraisal" are included. The automatic "appraisal" process is simple and usually fast, mainly involving implicit processes [7]. The deliberative appraisal process is more explicit and usually slower. These appraisal processes are carried out by a combination of the ACS, the NACS, the MCS, and the MS within CLARION. Among these subsystems, the NACS, as discussed in section "A Comprehensive Framework Capable of Addressing Emotion," is mainly responsible for reasoning (implicit or explicit), including when reasoning is carried out for deliberative appraisal. The MS, the MCS, and the ACS, especially their implicit processes, are responsible for the fast, automatic "appraisal" (gut reactions) as discussed above. We term this fast process "reactive affect," in order to distinguish it from deliberative appraisal. However, with reactive affect, the MCS or the ACS may trigger deliberative appraisal.

With the generation of emotion (based on different kinds of appraisal), there is the need for action or coping. Coping of emotion as identified by, for example, Lazarus and Folkman [19] can be carried out in CLARION through the ACS and the NACS. Among them, coping by the ACS is obviously action-oriented, but the actions may be either internally or externally oriented (as indicated in section "A Comprehensive Framework Capable of Addressing Emotion"), while coping by the NACS may be centered on reasoning (implicit or explicit) for the sake of, for example, re-evaluating the situations.

There are also other, more subtle types of emotion regulation (see e.g., [11]). In general, in CLARION, regulation of emotion may be accomplished in a number of ways at different phases of processing: for example, (1) at the perceptual phase (e.g., by preventing the perception of threatening stimuli), (2) at the motivational phase (e.g., by setting or changing priorities), (3) at the appraisal phase (e.g., by directing or re-directing appraisal), or (4) at the action phase (e.g., by suppressing or enabling certain types of actions). So emotion regulation can be carried out through suppression, enabling, re-evaluation, or other relevant means. It can be done either implicitly or explicitly (or in both ways). For instance, Scherer [39] identified *automatic unconscious regulation*, which involves (implicit) information filtering, task switching, or other generally rapid reactions, as well as *controlled conscious regulation*. There is a need to bridge emotion regulation and emotion generation. When emotion generation and appraisal are incorporated into detailed computational models, these models need to incorporate sophisticated regulatory mechanisms also (which are often absent in existing models, beside simple forms of coping).

Within CLARION, more sophisticated emotion regulation can be accomplished through the MCS (as its name suggests). The MCS regulates in response to sensory inputs, motivational states, and appraisal. Regulation by the MCS takes the form of input filtering, goal setting, action output filtering, and the like (see section "A Comprehensive Framework Capable of Addressing Emotion"; [45, 49]), corresponding to these phases identified earlier. Emotion regulation may thus directly affect action and reasoning within the ACS and the NACS. At a deeper level, drive activations within the MS may also be adjusted as a form of emotion regulation, through the MCS (e.g., by adjusting the gain parameters within the MS); thus action and reasoning (within the ACS and the NACS) change as a result.

In this regard, note that a clear distinction between emotion generation (e.g., through motivation, appraisal, and action) and emotion regulation (e.g., of inputs, of action outputs, of reasoning, and of motivation) is unnecessary.

A Model of Emotion within CLARION

From the discussions thus far, a mechanistic (computational) model of emotion is ready to emerge within the CLARION framework. Let us put the pieces together. Below, a conceptual-level model of emotion is outlined within the CLARION framework. Its mechanistic (computational) underpinning within the CLARION cognitive architecture is then sketched (for further technical details, see [51, 61]).

Outline of the Model

A general outline of our model of emotion is as follows. We divide emotion roughly into three aspects: reactive affect, deliberative appraisal, and coping/action, as identified in section "Addressing Issues of Emotion within the Framework" above (see also [61, 62]). These aspects will be discussed below one by one.

First, look into reactive affect. The notion of affect is generally poorly defined, and it tends to refer to many different things. However, in relation to emotion, in our model, we view affect as a fast, reactive component that precedes other, slower components of emotion, as alluded to earlier in section "Addressing Issues of Emotion within the Framework" [51, 67]. Emotion-evoking stimuli may impact an individual before (or without) conscious awareness. Empirical findings suggest that there is an independent mechanism that drives unconscious experiences of emotion [29, 65]. Processes surrounding unconscious emotional experiences are separate from slower, more deliberative processes. They are fast, reactive, and implicit (as is consistent with the CLARION framework). Reactive affect may be either positive or negative in valence, and to different degrees of intensity (i.e., arousal; [60]).

Next, look into deliberative appraisal, as identified in section "Addressing Issues of Emotion within the Framework." Such appraisal is the evaluation of the significance of an event or a situation [18, 39] on the basis of various inputs including reactive affect. Scherer et al. [40] postulated that emotion resulted from the assessment of a situation according to criteria such as *relevancy, implications, coping potential,* and *normative significance*. Other relevant dimensions of assessment that have been proposed include *cause, desirability, and likelihood*.[2] For instance, if Mary desires for her preferred candidate to win an election and believes that he/she will win, then she may experience the (potentially conscious) emotion of elation. Elation in this situation is associated with high values for "likelihood" and "desirability" (see section "Emotion and Motivation"). Emotions are shaped by cognitive evaluation (deliberative appraisal) of situations and events in our model. Deliberative appraisal also determines how emotions are consciously registered and reported (i.e., emotion terms/labels), and recommends possible goals and actions to pursue, according to our model.

In particular, emotion terms or labels are the outcomes of this process [40]. Much research has been focused on the structural relationship between appraisal variables and specific emotion labels. Various researchers have defined categories and sub-categories of these emotion labels [17, 43], although disagreement does exist.

Turn now to consider how outcomes of deliberative appraisal are applied. The evaluation of an event or situation may lead to specific action tendencies [40]. Existing theories suggest that outcomes of appraisal may lead to specific physiological, behavioral, or motivational changes, or to specific requirements for further information processing. The relationship between appraisal variables and various forms of responses has been explored (see, e.g., [24, 39]). Moreover, appraisal can trigger cognitive and behavioral responses, which in turn can become inputs into further appraisal, thus creating a continuous cycle of appraisal and reappraisal [17].

[2] Some of the appraisal dimensions might contain sub-dimensions. Some of the dimensions may also represent intermediate processing steps. For example, the "cause" dimension might require a process that can associate environmental factors with beliefs about the cause(s) of those factors.

Reactive affect and deliberative appraisal considered together, it seems reasonable to view appraisal in general, in a broad sense, as made up of both conscious (explicit) and subconscious (implicit) processes that use various sources of inputs to generate beliefs, goals, affect, and emotion labels, which can subsequently guide action decision making as well as regulatory processes (as identified in section "Addressing Issues of Emotion within the Framework"; more later).

Third, we examine coping and action. Coping is defined as cognitive and behavioral efforts at managing specific external and/or internal demands. Coping often follows appraisal [17, 19]. In order to cope with a situation, one must ascertain the meanings of the situation, which include one's reactive affect about a situation, what one believes the likely outcome of that situation may be (e.g., from deliberative appraisal), and so on. Once meanings are ascertained, selection of a coping strategy is made and a pattern of behavior initiated.

A fundamental reason why behaviors are performed is the pursuit of basic needs [15], but a direct link between basic needs and specific behaviors may not be realistic. Needs may be implicit and thus inaccessible [15, 25, 46, 56]. More specific mental constructs are needed to capture the means by which needs are attended to. These constructs are goals: that is, *a well-defined target state that is actively pursued* [31]. Thus, the link between appraisal and behavior is established via goals (as posited in CLARION, discussed in section "A Comprehensive Framework Capable of Addressing Emotion"; see [46, 49]).

Goals can be used to capture more internally oriented coping (a form of self-regulation), in addition to purely externally oriented behaviors. Using an election example, a goal to *prevent one's preferred candidate from losing the election* may lead to more problem-focused (externally oriented) behaviors (e.g., *give money* or *volunteer*), while the goal to *reduce the effects of the anxiety-inducing stimulus* may result in more emotion-focused (internally oriented) behaviors (e.g., *turn off the TV* or *stop thinking about it*). Goals may also lead to re-appraisal (as discussed earlier).

Considered in total, according to CLARION, behaviors are determined by goals, which are selected based on input state information, drives, reactive affect, deliberative appraisal, and other factors. Goal setting orients an individual toward certain types of actions. Coping consists of both internally and externally oriented actions. Beyond that, there may also be other forms of self-regulation (e.g., aimed at altering motivation or affect, or at filtering perception).

CLARION Specifics

Let us explore further these three aspects within the CLARION computational cognitive architecture specifically. In a nutshell, according to the cognitive architecture, both reactive affect and coping/action are captured by the dynamics among the mechanisms of the MS, the ACS, and the MCS, and deliberative appraisal is captured by the mechanisms within the NACS.

First, look into reactive affect. In CLARION, reactive affect may be determined based on motivation as well as potentials for action [61]. According to the literature, it may have a lot to do with the relationship between situations/events and an individual's desires and intentions (i.e., drives and goals; [34]). The MS of CLARION contains drives (at the bottom level) and goals (at the top level) and collectively captures the processes by which an individual is compelled to act, as outlined in section "A Comprehensive Framework Capable of Addressing Emotion" [46]. Therefore, the MS can appropriately capture the role of motivation in reactive affect.[3]

On the other hand, within CLARION, the bottom (implicit) level of the ACS contains neural networks (see section "A Comprehensive Framework Capable of Addressing Emotion"; [44, 49]). These networks propagate activation from input nodes representing input state information to output nodes representing actions.[4] Before the final action decision is made, the activations of the output nodes really just represent potentials to act, or "action potentials." These activations capture the degree and likelihood that the actions will lead to a desirable outcome (see section "A Comprehensive Framework Capable of Addressing Emotion" regarding Q values). Fundamentally, actions are chosen based on their ability to satisfy needs [25]. Thus, the "action potential" represents the expected degree to which a set of actions (that can be started in the current state) will be successful in attending to the needs of an individual.[5]

Within the CLARION framework, we may look at various combinations of drive activation and action potential and relate these combinations to experiences of positive and negative reactive affect:

- When some drives are highly activated (i.e., some needs must be attended to), but action potentials are also high (i.e., those needs are likely to be met), then affect could not be very negative—it could range from highly positive to slightly negative. For instance, if the highly activated is an avoidance-oriented drive (e.g.,

[3] The MS has been justified extensively elsewhere [46, 49]. It has been used to capture many motivationally based phenomena (e.g., [50, 52, 63]).

[4] The ACS may recommend actions using a combination of the top (explicit) and the bottom (implicit) level, but the bottom level is always activated as part of the decision-making process [44, 49].

[5] Note that for calculating action potentials, inputs to the neural networks might include drive activations but not goals (because it might be done before goal setting).

"avoiding danger"; see section "A Comprehensive Framework Capable of Addressing Emotion"), affect should be only slightly negative (which may lead to, e.g., slight anxiety; see section "Emotion and Motivation"). If the highly activated is an approach-oriented drive (e.g., "recognition and achievement"), affect should be positive (which may lead to, e.g., elation; see section "Emotion and Motivation").
- When some drives are highly activated (i.e., some needs are high) and action potentials are low (i.e., those needs will likely not be met), affect could range from slightly negative to very negative. For instance, if some avoidance-oriented drives (e.g., "avoiding danger") are highly activated in this case, affect should be very negative (leading likely to high anxiety or fear).
- When drive activations are low (i.e., needs are being met) and action potentials are high (i.e., needs will likely continue to be met), affect should be somewhere in the positive range.
- When drive activations are low (i.e., needs are being met) and action potentials are low (i.e., needs may not continue to be met), affect should not be very positive (but should not be very negative either).

The exact magnitude of the affect in each of these circumstances above is drive specific (and may also be context dependent to some degree). Thus, they need to be specified on a case-by-case basis. (They will not be enumerated here; but see, e.g., [61].)

Reactive affect generated in this way can influence certain implicit processes. For instance, one regulatory process is the orientation toward certain types of motivation, approach- or avoidance-oriented, on the basis of either positive or negative affect [49]. This adjustment is carried out within CLARION through adjusting the gain parameters (i.e., $gain_s$; see section "A Comprehensive Framework Capable of Addressing Emotion"). In addition, a positive affective state can serve as a positive stimulus and a negative affective state as a negative stimulus (by treating the reactive affect as part of the input state). Beyond such fast, reactive and implicit application, reactive affect is also an important contributor to the slower deliberative appraisal.

In CLARION, the NACS is for storing and utilizing declarative knowledge in various forms [49]. It contains mechanisms by which knowledge can be compared, associated, and otherwise reasoned over [12, 49].[6] As demonstrated in Wilson [61], the NACS provides the representational and mechanistic means by which various aspects of deliberative appraisal may be actualized.

[6] Previous research has shown that the NACS captures many aspects of human reasoning, including similarity-based reasoning, rule-based reasoning, analogical reasoning, incubation, insight, and creativity [12, 53].

For example, the following explicit appraisal dimensions (as discussed earlier) may be assessed through reasoning within the NACS:

- *Relevance*—the direction (i.e., being positive or negative) and the intensity of the situation (in part based on reactive affect, but also based on declarative knowledge)
- *Implications*—the likelihood, unexpectedness, and changeability of the situation as well as goal congruence (in part based on action potentials)
- *Coping potential*—the subjective sense of control over the event (in part based on action potentials)
- *Normative significance*—the attribution of cause to the situation

Specifically, these appraisal dimensions, as well as goals and emotion terms, are represented in CLARION as *chunks* (as described in section "A Comprehensive Framework Capable of Addressing Emotion"). To assess these, prior knowledge and beliefs are used within the NACS (when triggered by current situations). Chunks can be naturally inferred using the reasoning processes existing in the NACS at the top level and/or the bottom level (see section "A Comprehensive Framework Capable of Addressing Emotion," as well as [49]). Beliefs concerning appraisal dimensions, goals, and emotion terms can be inferred in the form of chunk activation. As shown by Wilson [61], the NACS reasoning mechanisms provide the necessary medium for deliberative appraisal, and capture several existing theories of appraisal.

The outcomes of deliberative appraisal may affect motivation in two different ways: (1) the outcomes may be used to decide on an appropriate goal to satisfy activated drives (needs), which can be done by the MCS (or the ACS); (2) the outcomes of appraisal may affect drive activations (e.g., through treating the outcomes as part of the input state, or through adjusting the approach/avoidance orientation by the MCS). The outcomes of deliberative appraisal may also affect action decision making within the ACS through treating the outcomes as part of the input state for the ACS.[7]

Consciously recognized emotion is the outcome of deliberative appraisal. For instance, as discussed earlier, it has been hypothesized within CLARION that the emotion of *elation* is related to positive reward or "expectation" of positive reward. Specifically, first, the corresponding reactive affect may be, computationally, the result of high

[7] The outcomes of deliberative appraisal from the NACS may be filtered by the MCS. This allows the MCS to select only the knowledge that is relevant to the situation. Filtering could be done based on the current input state and its relevant microfeatures (see section "A Comprehensive Framework Capable of Addressing Emotion").

activation of an approach-oriented drive and a high action potential, which indicates the expectation of related reward (or the result of high activation of an approach-oriented drive and high reward). Deliberative appraisal occurring within the NACS then assesses the situation and produces a label for the felt affect, among other things.

For another instance, it has also been hypothesized that the emotion of *anxiety* can be related to "expectation" of negative reward. Computationally, the corresponding reactive affect is the result of high activation of an avoidance-oriented drive and a relatively low action potential, which indicates a low probability of avoiding negative reward. The intensity of anxiety is a function of avoidance-oriented drive activations and action potentials. Compared with activation of an approach-oriented drive, here the role of the action potential is less significant—an avoidance-oriented drive arouses a certain degree of anxiety or fear regardless of levels of action potentials. Deliberative appraisal occurring within the NACS then assesses the situation and produces a label for the felt affect.

Coping/action follows. As outlined earlier, results from deliberative appraisal (represented by chunks) may be applied to (1) directly impact action decision making as inputs, (2) set goals to initiate behavior through either externally oriented (i.e., problem-focused) or internally oriented (i.e., emotion-focused) goals, and (3) initiate regulatory processes. The MCS captures many of such regulatory functions.[8]

In particular, action decision making within the ACS may be facilitated by the outcomes of deliberative appraisal. Just like declarative knowledge in general, chunks concluded by the NACS may becomes available as inputs to the ACS for action decision making (see section "A Comprehensive Framework Capable of Addressing Emotion").

Goals are important with regard to coping. Recommended goals may be generated from deliberative appraisal by processes within the NACS. Goals can then be set by the MCS using a context-dependent method. Specifically, in the MCS, to take account of deliberative appraisal in setting goals, goal strengths are determined by both drive activations and the outcomes of deliberative appraisal. In other words, goals are set based not only on their relevance (as in the goal strength equation of section "A Comprehensive Framework Capable of Addressing Emotion"), but also on appraisal regarding their feasibility and benefits to the individual (see [61] for the extended equation).[9]

Once goals are selected, those behaviors that best facilitate goal achievement are likely to be selected by the ACS.[10] The action selection mechanisms are well defined in CLARION (as briefly reviewed in section "A Comprehensive Framework Capable of Addressing Emotion" earlier).

Simulations, Comparisons, and Summary

Several simulations have been carried out within CLARION. For instance, Wilson and Sun [62] show how CLARION may be used to capture the emotional dynamics of victims of school bullying. It provides a unique mechanistic interpretation of the appraisal process and the coping strategy selection by the victims of school bullying. It also demonstrates how such processes can be precisely expressed computationally and therefore simulated within CLARION. A number of other examples were also dealt with (see, e.g., [61–64]).

In the past, a variety of computational models of emotion were proposed in the literature. Existing computational models of emotion include: [14, 21, 24, 28, 34, 38], and many others. Some of them bear some similarities to the present model. For example, [23] integrated cognition, emotion, and learning to some extent in an AI model. Marsella and Gratch [24] instantiated their model of appraisal and coping within Soar (although not psychologically validated). Bach [1] did include motivation in addressing emotion (but the model was not extensively validated against psychological data).

Some distinguishing features of the present work include: (1) The present model is based on the foundation of a generic but detailed, psychologically realistic cognitive architecture (which has undergone development and psychological validation for over two decades based on a wide variety of psychological data), and thus based on a more comprehensive and better grounded view of the architecture of the mind. (2) The same cognitive architecture has led to unified explanations of a wide range of psychological phenomena, in addition to and together with emotional phenomena; hence it is broad and integrative. (3) The present model includes a motivational subsystem that synthesized well developed theories on motivational processes (e.g., [30, 35, 46, 47]), which is not commonly found among psychologically validated cognitive models. (4) The present model captures emotion, in a unique way, based on drives, goals, and actions, whereby drives serve as

[8] In CLARION, the MCS is responsible for many regulatory processes [49], including, among others, goal setting, parameter changing, and input and output filtering. In previous work, the MCS has been shown to capture a variety of psychological phenomena (see, e.g., [63]).

[9] Prior research has extensively explored goal setting using the MCS [46, 52].

[10] These CLARION details outlined above address the processes by which coping strategies are chosen. They also suggest a possible origin for how the appraisal dimensions might be formed, that is, possibly for the sake of facilitating goal setting.

the basis for setting goals and consequently for reasoning and for action, along with metacognitive regulation.

In CLARION, emotion involves a range of subsystems: the ACS (for action), the NACS (for reasoning), the MS (for motivation), and the MCS (for metacognitive regulation). Complex interactions occur among these different subsystems and among many components within. So the present model is a complex dynamic system. But it is a complex dynamic system with clearly structured components (each with specific knowledge, mechanisms, and processes) interacting with each other.

This overview article has so far shown that, although admittedly still preliminary, CLARION has the potential for providing unique, broad, and integrative mechanistic interpretations of issues surrounding emotion. Exploring emotion within a comprehensive computational cognitive architecture enables its theorizing and modeling to make contact with detailed, established psychological mechanisms and processes. As a result, the study of emotion is naturally and seamlessly linked to other psychological processes and mechanisms such as memory, decision making, reasoning, motivation, and metacognitive regulation, defined generically within a cognitive architecture [51, 61]. The result is a broad, psychologically realistic model that provides psychologically well-grounded interpretations. The present article has tried to provide an overview of this project, pulling all the (previously devised and tested) pieces together.

Acknowledgments This work has been supported in part by the ONR Grants N00014-08-1-0068 and N00014-13-1-0342.

Compliance with Ethical Standards

Conflict of Interest Ron Sun, Nick Wilson, and Michael Lynch declare that they have no conflict of interest.

Informed Consent All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008 (5). Additional informed consent was obtained from all patients for which identifying information is included in this article.

Human and Animal Rights This article does not contain any studies with human or animal subjects performed by any of the authors.

References

1. Bach J. Principles of synthetic intelligence PSI: an architecture of motivated cognition. New York: Oxford University Press; 2009.
2. Bower GH. Mood and memory. Am Psychol. 1981;36:129–48.
3. Carver C, Scheier M. On the self-regulation of behavior. Cambridge: Cambridge University Press; 1998.
4. Clark LA, Watson D. Temperament: a new paradigm for trait psychology. In: Pervin LA, John OP, editors. Handbook of personality: theory and research. 2nd ed. New York: Guilford Press; 1999. p. 399–423.
5. Damasio A. Descartes' error: emotion, reason, and the human brain. New York: Penguin; 1994.
6. Dorner D. The mathematics of emotion. In: Proceedings of the fifth international conference on cognitive modeling. Germany: Bamberg; 2003. p. 75–79.
7. Ekman P. Basic emotions. In: Dalgleish T, Power M, editors. Handbook of cognition and emotion. Sussex: Wiley; 1999.
8. Evans J, Frankish K, editors. In two minds: dual processes and beyond. Oxford: Oxford University Press; 2009.
9. Flavell J. Metacognitive aspects of problem solving. In: Resnick B, editor. The nature of intelligence. Hillsdale: Erlbaum; 1976.
10. Frijda N. The emotion. Cambridge: Cambridge University Press; 1986.
11. Gross JJ, editor. Handbook of emotion regulation. New York: Guilford Press; 2007.
12. Helie S, Sun R. Incubation, insight, and creative problem solving: a unified theory and a connectionist model. Psychol Rev. 2010;117(3):994–1024.
13. Higgins ET. Beyond pleasure and pain. Am Psychol. 1997; 52(12):1280–300.
14. Hudlicka E. A computational model of emotion and personality: applications to psychotherapy research and practice. In: Proceedings of the 10th annual cybertherapy conference: a decade of virtual reality. Switzerland: Basel; 2005.
15. Hull C. Essentials of behavior. New Haven: Yale University Press; 1951.
16. Lambie J, Marcel A. Consciousness and the variety of emotion experience: a theoretical framework. Psychol Rev. 2002;109(2): 219–59.
17. Lazarus RS. Cognition and motivation in emotion. Am Psychol. 1991;46(4):352–67.
18. Lazarus RS. Relational meaning and discrete emotions. In: Scherer KR, Schorr A, Johnstone T, editors. Appraisal processes in emotion. Oxford: Oxford University Press; 2001.
19. Lazarus RS, Folkman S. Stress, appraisal, and coping. New York: Springer; 1984.
20. Ledoux J. Brain mechanisms of emotion and emotional learning. Curr Opin Neurobiol. 1992;2(2):191–7.
21. Leven S, Levine D. Multiattribute decision making in context: a dynamic neural network methodology. Cogn Sci. 1996;20: 271–99.
22. Lewis M. Bridging emotion theory and neurobiology through dynamic systems modeling. Behav Brain Sci. 2004.
23. Marinier RP III, Laird JE, Lewis RL. A computational unification of cognitive behavior and emotion. Cogn Syst Res. 2009;10(1): 48–69.
24. Marsella S, Gratch J. EMA: a process model of appraisal dynamics. Cogn Syst Res. 2009;10(1):70–90.
25. Maslow A. Motivation and personality. 3rd ed. New York: Harper and Row; 1987.
26. Miceli M, Castelfranchi C. Expectancy and emotion. Oxford: Oxford University Press; 2014.
27. Mineka S, Sutton S. Cognitive biases and the emotional disorders. Psychol Sci. 1992;3(1):65–9.
28. Mueller ET, Dyer MG. Daydreaming in humans and computers. In: Proceedings of the ninth international joint conference on artificial intelligence. Los Angeles: Morgan Kaufmann Publishers, Inc.; 1985. p. 18–24.
29. Murphy ST, Zajonc RB. Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures. J Pers Soc Psychol. 1993;64:723–39.
30. Murray H. Explorations in personality. New York: Oxford University Press; 1938.

31. Ortony A, Clore G, Collins A. The cognitive structures of emotions. New York: Cambridge University Press; 1988.
32. Reber A. Implicit learning and tacit knowledge. J Exp Psychol Gen. 1989;118(3):219–35.
33. Reder L, editor. Implicit memory and metacognition. Mahwah: Erlbaum; 1996.
34. Reisenzein R. Emotions as metarepresentational states of mind: naturalizing the belief-desire theory of emotion. Cogn Syst Res. 2009;10(1):6–20.
35. Reiss S. Multifaceted nature of intrinsic motivation: the theory of 16 basic desires. Rev Gen Psychol. 2004;8(3):179–93.
36. Rumelhart D, McClelland J, PDP Research Group. Parallel distributed processing: explorations in the microstructures of cognition. Cambridge: MIT Press; 1986.
37. Schacter D. Implicit memory: history and current status. J Exp Psychol Learn Mem Cogn. 1987;13:501–18.
38. Scherer K. Studying the emotion-antecedent appraisal process: the expert system approach. Cogn Emot. 1993;7:325–55.
39. Scherer KR. Emotion and emotional competence: conceptual and theoretical issues for modelling agents. In: Scherer K, Banziger T, Roesch E, editors. A blueprint for affective computing. New York: Oxford University Press; 2010. p. 3–20.
40. Scherer KR, Schorr A, Johnstone T, editors. Appraisal processes in emotion. Oxford: Oxford University Press; 2001.
41. Simon HA. Motivational and emotional controls of cognition. Psychol Rev. 1967;74:29–39.
42. Smillie LD, Pickering AD, Jackson CJ. The new reinforcement sensitivity theory: implications for personality measurement. Pers Soc Psychol Rev. 2006;10:320–35.
43. Smith CA, Lazarus RS. Emotion and adaptation. In: Pervin LA, editor. Handbook on personality: theory and research. New York: Guilford; 1990. p. 609–37.
44. Sun R. Duality of the mind. Mahwah: Lawrence Erlbaum Associates; 2002.
45. Sun R. A tutorial on CLARION 5.0. Technical report, Cognitive Sciences Department, Rensselaer Polytechnic Institute, Troy. 2003. http://www.cogsci.rpi.edu/~rsun/sun.tutorial.pdf.
46. Sun R. Motivational representations within a computational cognitive architecture. Cogn Comput. 2009;1(1):91–103.
47. Sun R. Theoretical status of computational cognitive modeling. Cogn Syst Res. 2009;10(2):124–40.
48. Sun R. Memory systems within a cognitive architecture. New Ideas Psychol. 2012;30:227–40.
49. Sun R. Anatomy of the mind. New York: Oxford University Press; 2016.
50. Sun R, Fleischer P. A cognitive social simulation of tribal survival strategies: the importance of cognitive and motivational factors. J Cogn Cult. 2012;12(3–4):287–321.
51. Sun, R, Mathews R. Implicit cognition, emotion, and meta-cognitive control. Mind Soc. Special issue on Dual Processes Theories of Thinking (edited by D. Over, L. Macchi, and R. Viale), 2012;11(1):107–119.
52. Sun R, Wilson N. A model of personality should be a cognitive architecture itself. Cogn Syst Res. 2014;29–30:1–30.
53. Sun R, Zhang X. Accounting for a variety of reasoning data within a cognitive architecture. J Exp Theor Artif Intell. 2006;18(2):169–91.
54. Sun R, Merrill E, Peterson T. From implicit skills to explicit knowledge: a bottom-up model of skill learning. Cogn Sci. 2001;25(2):203–44.
55. Sun R, Slusarz P, Terry C. The interaction of the explicit and the implicit in skill learning: a dual-process approach. Psychol Rev. 2005;112(1):159–92.
56. Toates F. Motivational systems. Cambridge: Cambridge University Press; 1986.
57. Tolman EC. Purposive behavior in animals and men. New York: Century; 1932.
58. Tyrell T. Computational mechanisms for action selection. Ph.D. Thesis, Oxford: Oxford University; 1993.
59. Watkins C. Learning with delayed rewards. Ph.D. Thesis, Cambridge: Cambridge University; 1989.
60. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol. 1988;54:1063–79.
61. Wilson NR. Towards a psychologically plausible comprehensive computational theory of emotion. Ph.D Thesis, Cognitive Sciences Department, Troy: Rensselaer Polytechnic Institute; (2012).
62. Wilson N, Sun R. Coping with bullying: a computational emotion-theoretic account. In: Proceedings of the annual conference of cognitive science society, Quebec City, Quebec, Canada. p. 3119–24. Published by Cognitive Science Society, Austin, Texas (2014).
63. Wilson N, Sun R, Mathews R. A motivationally-based simulation of performance degradation under pressure. Neural Netw. 2009;22:502–8.
64. Wilson N, Sun R, Mathews R. A motivationally based computational interpretation of social anxiety induced stereotype bias. In: Proceedings of the annual conference of cognitive science society, Oregon: Portland. p. 1750–55; 2010.
65. Winkielman P, Berridge K, Wilbarger J. Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. Pers Soc Psychol Bull. 2005;31(1):121–35.
66. Wright IP, Sloman A. MINDER1: an implementation of a protoemotional agent architecture. Technical Report CSRP-97-1, University of Birmingham, School of Computer Science; 1997.
67. Zajonc RB. Feeling and thinking: preferences need no inference. Am Psychol. 1980;35(2):151.