# LIDA: A Systems-level Architecture for Cognition, Emotion, and Learning

Stan Franklin, Tamas Madl, Sidney D'Mello and Javier Snaider

*Abstract*—**We describe a cognitive architecture (LIDA) that affords attention, action selection and human-like learning intended for use in controlling cognitive agents that replicate human experiments as well as performing real-world tasks. LIDA combines sophisticated action selection, motivation via emotions, a centrally important attention mechanism, and multimodal instructionalist and selectionist learning. Empirically grounded in cognitive science and cognitive neuroscience, the LIDA architecture employs a variety of modules and processes, each with its own effective representations and algorithms. LIDA has much to say about motivation, emotion, attention, and autonomous learning in cognitive agents. In this paper we summarize the LIDA model together with its resulting agent architecture, describe its computational implementation, and discuss results of simulations that replicate known experimental data. We also discuss some of LIDA's conceptual modules, propose non-linear dynamics as a bridge between LIDA's modules and processes and the underlying neuroscience, and point out some of the differences between LIDA and other cognitive architectures. Finally, we discuss how LIDA addresses some of the open issues in cognitive architecture research.**

*Index Terms*—**Autonomous agent, Cognitive model, Computational model, Cognitive architecture, LIDA, Agent architecture, Perceptual learning, Episodic learning, Procedural learning, Action-perception cycle, Cognitive cycle, Neural correlates, Affordance, Attention, Action selection, Emotions**

## I. Introduction

As social psychologist Kurt Lewin so succinctly pointed out "There is nothing so practical as a good theory" [1, p. 169]. Artificial intelligence pioneer Allen Newell strongly supported the need for systems-level theories/architectures, asserting that "You can't play 20 questions with nature and win" [2]. More recently, memory researcher Douglas Hintzman, echoing Newell in decrying the reliance on modeling individual laboratory tasks, stated that "Theories that parsimoniously explain data from single tasks will never generalize to memory as a whole…" [3]. Cognitive architects Langley, Laird and Rogers argue that "Instead of carrying out micro-studies that address only one issue at a time, we should attempt to unify many findings into a single theoretical framework, then proceed to test and refine that theory" [4]. In line with these views, this paper presents a summary account of our systems-level conceptual LIDA cognitive model ($\mathrm{LIDA_C}$) together with its implemented computational cognitive architecture ($\mathrm{LIDA_I}$) as a candidate for the unified theoretical framework called for above. Discussing LIDA's contributions to the open issues in cognitive modeling listed by Langley et al, [4], as well as its answers to previously suggested criteria for models of human cognition [5], we argue that LIDA is a plausible candidate for a unified theoretical framework of cognition.

The fundamental principle guiding LIDA is that every autonomous agent [6], be it human, animal or artificial (e.g., software agent, robot), must frequently and continually sense its environment, interpret what it senses, and then act. Ecological psychologists and cognitive neuroscientists refer to this as the action-perception cycle [7, 8]. An agent must select appropriate actions to further its goals, depending on affordances in its environment. Thus, action selection is central to cognition (see the action selection paradigm [9]) and is the overriding task of every broad cognitive architecture, e.g., SOAR [10], ACT-R [11], CLARION [12], etc.

For more sophisticated cognitive agents, action selection requires the agent to understand its current circumstances (i.e. the context), that is, the frequently recurrent transformation of sensory data into an internal model of its current situation. Many such cognitive agents gain from (or are burdened with) multiple sensory systems that produce huge amounts of data from their complex, dynamic environments. There is often too much information to attend to at once. A cognitive architecture controlling one of these agents would benefit from some mechanism for attention [13, 14] that would choose the most salient portion of the current situation for the agent to attend to while selecting its next action. Indeed, it has been argued that attentional mechanisms are vital for handling real-world complexity, since the number of combinations of memory items, percepts, and possible actions can be extremely high, but agents have limited resources for selecting a suitable action [4, 15]. However, explicit general attentional mechanisms are not commonly included in cognitive architectures [4], although some architectures model some aspects of attention, such as visual attention [16, 17], eye-movements [18, 19], and multi-tasking [20, 21]. Attention in

Sidney D'Mello is supported by the National Science Foundation (NSF) (ITR 0325428, HCC 0834847, DRL 1235958) and Institute of Education Sciences (IES) through Grant R305A080594. Tamas Madl is supported by EPSRC (Engineering and Physical Sciences Research Council) grant EP/I028099/1. Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF, IES, DoE, or EPSRC.

Stan Franklin is with the University of Memphis, Memphis, TN 38152 USA (phone: (901) 678-1341; fax: (901) 678-5129; e-mail: franklin@ memphis.edu).

Tamas Madl is with the University of Manchester, UK. (e-mail: tamas.madl@gmail.com).

Sidney D'Mello is with the University of Notre Dame, Notre Dame, IN 46556 USA (e-mail: sdmello@nd.edu).

Javier Snaider is with the University of Memphis, Memphis, TN 38152 USA (e-mail: jsnaider@memphis.edu).

LIDA is the process of bringing content to consciousness, following [22]. It is general in the sense that it is independent from modality, can theoretically focus on perceived, remembered, or imagined content, and is thus able to model a wide variety of paradigms using the same mechanism (see the $LIDA_I$ Attention [23] and Attentional Blink [24] agents in Section X).

Handcrafting sophisticated cognitive agents “living” in complex environments is often prohibitively expensive, if not impossible. Thus, many of the cognitive architectures employ some form of learning (e.g. [10, 12]). Some of these learning mechanisms suffer from being either unimodal or supervised. Supervised learning has two major drawbacks. Often, a large and expensive training set, which is sometimes unavailable, is required. Also, the designer must know in advance what sort of representations must be learned. Unimodal learning, mostly procedural (i.e., learning of new productions), often ignores other critical aspects of the agent’s cognitive processes, such as perception. Cognitive architectures with more human-like, unsupervised, and multimodal learning mechanisms are needed to permit continual, effective selectionist (reinforcement) and instructionalist (new representations) learning by a cognitive agent.

The LIDA[1] model, described below, provides an example of a cognitive architecture combining sophisticated action selection, motivation and learning, a centrally important attention mechanism, and multimodal instructionalist and selectionist learning [25]. Empirically grounded in cognitive neuroscience, the LIDA architecture is neither symbolic nor connectionist, but blends crucial features of each. It employs a variety of modules[2] and processes, each with its own effective representations and algorithms. Most of these involve multiple, interactive, memory systems.

Every animal with a nervous system learns about the structure of its world by encoding entries into memory systems [26]. The later retrieval of some of these memories facilitates adaptive responses to a changing environment. Thus, every systems-level cognitive model used for studying human and animal minds, as well as artificial minds, must concern itself with learning and memory. Memory can be partitioned into multiple, interactive memory systems in various ways useful for different purposes [27-29]. In the LIDA model we categorize memory according to the type of structures learned, its function, and its duration. The result is a bevy of interactive memory systems: sensory, perceptual (recognition), spatial, episodic (two varieties), attentional, procedural, and sensory-motor, as depicted in Figure 1. Memory and learning play a central role in the LIDA model. Learning is mostly conceptual as of yet, due to the only recently released $LIDA_I$ computational framework. That being said, $LIDA_I$ implementations exist for most memory modules, and for some learning mechanisms (see Sections VI and XI).

In order to distinguish the conceptual ideas of the LIDA model from the mechanisms that have already been implemented computationally, we will use the following subscript notation: $LIDA_C$ for the conceptual model, $LIDA_I$ for the implemented parts, and LIDA without a subscript to refer to both. For example, “LIDA accounts for functional consciousness” implies that a functional consciousness mechanism is part of both the conceptual and the computational model (although strongly simplified in the latter); whereas “$LIDA_C$ can model feelings” means that feelings are part of the conceptual but not the computational model.

Though aspects of the LIDA model have previously been described in several short papers, [30-38], this paper provides a summative account by integrating the various components of the model. Additional contributions include: (a) a discussion of the design principles underlying the LIDA model, (b) a brief description of the software implementation of the $LIDA_I$ Framework, (c) a short account of four $LIDA_I$-based software agents that replicate experimental results, (d) a more detailed description of $LIDA_C$’s use of feelings and emotions as motivators, (e) a discussion of the relationship of the $LIDA_C$ conceptual model and the underlying neural architecture of brains, and (f) a comparison between LIDA and some of the more popular cognitive architectures.

It is imperative to note that LIDA is not intended to model brains. That is, as a cognitive model, the LIDA model is concerned with the functional organization of cognitive processes at the conceptual and computational level at which they are studied in cognitive science. Though the model must be empirically validated by both psychological and neuroscientific studies, it makes no attempt to model the underlying neural anatomy or mechanisms. That does not mean, however, that the model is not inspired by neural mechanisms that are known, such as in visual or spatial cognition.

The LIDA model is partly computational and partly conceptual. This paper is organized to reflect this distinction. The first part of the paper describes the underlying theory with an emphasis on the aspects that have been computationally implemented and tested (Sections II, III, and IV). Section II explores the relationship between the LIDA model and other psychological and neuroscientific theories; Section III focuses on LIDA’s cognitive cycles, the high-level, but very brief processes out of which we hypothesize all cognition is composed. Section IV is devoted to decision making and action selection.

The next three sections provide brief discussions about some of the conceptual aspects of LIDA that have been designed, partially implemented, but not yet systematically tested. Section V describes how LIDA handles higher-level, multicyclic, cognitive processes. Section VI introduces LIDA’s ideas about learning, a centrally important part of the model. Section VII describes the use of feelings and emotions as motivators and facilitators of learning in the LIDA model. Further, Section VIII discusses how the high-level conceptual

[1] LIDA is an acronym for Learning Intelligent Distribution Agent (Learning IDA), where IDA is a software personnel agent hand-crafted for the US Navy that automates the process of finding new billets (jobs) for sailors at the end of a tour of duty. LIDA adds learning to IDA and extend its architecture in many other ways

[2] While the LIDA model seems modular, it makes no commitment to modularity in the underlying neural mechanisms.

LIDA model can be grounded in cognitive neuroscience, and tentatively outlines functional correspondences to neural correlates.

We then turn to concrete implementations of LIDA architecture via a description of the $LIDA_I$ computational framework (Section IX) and descriptions of four studies that use the computational framework to replicate human experimental data (Section X).

The last few sections focus on some of the broader issues stemming from the LIDA architecture. Section XI places LIDA in the context of other cognitive architectures, while Section XII discusses how LIDA might address some of the open issues in cognitive architectures that were raised in a review article by Langley et al. [4]. Finally, section XIII contains a brief conclusion.

## II. LIDA Model, Its Relations, and Design Principles

The LIDA model is a conceptual and partially implemented computational model that makes an attempt to cover a large portion of human cognition. It is largely based on Baars' Global Workspace Theory (GWT) [22, 39, 40], a conceptual theory of the role of consciousness[3] (specifically the attentional component[4]) in cognition. Originally conceived as a neuropsychological model of conscious and unconscious processes [22], GWT has been broadened in LIDA into a higher-level theory of human cognitive processing [31]. Now supported by considerable empirical evidence [40], GWT views the nervous system as a distributed parallel system with many different specialized processes. Coalitions of these processes enable an agent to make sense of the sensory data coming from the current environmental situation. Other coalitions, filtering and using this understanding, compete for attention in what Baars calls the global workspace. The contents of the winning coalition are broadcast globally, and are proposed to be phenomenally conscious. This conscious broadcast serves to recruit other unconscious processes to be used to select an appropriate response to the current situation. GWT is therefore a theory of how consciousness functions within cognition. The broadcast must be global to allow simultaneous learning into multiple memories with diverse functions.

This description of GWT is from the point of view of what happens during a single LIDA cognitive cycle (see the subsection Cognitive Cycles below). Viewing its contents over multiple successive cognitive cycles, the global workspace can be thought of as a fleeting memory system that enables access between brain functions that are otherwise separate (Baars, 2002). From this view it seems to be *"… a theater of mental functioning. Consciousness in this metaphor resembles a bright spot on the stage of immediate memory, directed there by a spotlight of attention under executive guidance. Only the bright spot is conscious, while the rest of the theater is dark and unconscious."* The hypothesized primary functional purpose of consciousness is to integrate, provide access, and coordinate the functioning of very large numbers of specialized networks that otherwise operate autonomously and unconsciously [41].

Besides GWT, the LIDA model implements and fleshes out a number of psychological and neuropsychological theories, including situated and grounded cognition [42, 43], perceptual symbol systems [42, 44], working memory [45, 46], memory by affordances [47], long-term working memory [48], and Sloman's H-CogAff cognitive framework [49]. This includes a broad array of cognitive modules and processes (discussed in Section III).

The LIDA computational architecture, derived from the LIDA cognitive model, employs a variety of modules that are designed using quite distinct computational mechanisms drawn from AI. These include variants of the Copycat Architecture [50, 51], Sparse Distributed Memory (SDM) [52, 53], the Schema Mechanism [54, 55], the Behavior Net [56, 57], and the Subsumption Architecture [58].

Please note that whenever we mention that our model accounts for some mental phenomenon, and use terms from cognitive science, we do not mean to imply that $LIDA_I$ is able to account for the full psychological complexity underlying these terms. Rather, we mean to say that these mental phenomena fit into and are part of the $LIDA_C$ model. If implemented as part of $LIDA_I$, their computational counterparts are functionally similar but very simple abstractions as is the case with most computational models.

It is important to emphasize some design principles that underlie the LIDA model (but are not necessarily unique to LIDA – see Section XI for comparisons with other cognitive architectures). Six such principles are discussed below. The first four principles have been implemented in the computational LIDA framework, while the last two are still conceptual.

**Principles of grounded cognition.** First, the model adheres to the *principles of grounded cognition* [42], which emphasize the importance of modal representations, situated action, and perceptual simulation. Instead of representing knowledge as amodal symbols in a semantic memory, the representations in the model, which resemble perceptual symbols [42, 44], are grounded in primitive sensors and actuators (see the Understanding phase in Section III, and vector representations in Section IX). Current $LIDA_I$ agents are not physically embodied[5], but interact with simulated environments, which can still implement the structural coupling between agent and environment which embodiment requires [59]. LIDA's predecessor, IDA, a software agent operating in a real-world virtual environment that included unstructured email correspondence with humans, was claimed to be embodied in this restricted sense [60]. There are now a large number of

[3] The LIDA model treats of functional consciousness, that is, consciousness as described in GWT (referring to information that is "broadcast" in the global workspace and made available to cognitive processes such as action selection, as opposed to only locally available, non-conscious information). It makes no commitment to phenomenal (subjective) consciousness.

[4] Following Baars (1988, p369), we think of attention as the process of bringing content to consciousness.

[5] Work is underway to physically embody LIDA on a PR2 humanoid robot

theoretical arguments [44, 61] as well as empirical findings [62, 63] (and many others - see [42] for a detailed review) in favor of a grounded as opposed to a symbolic or cognitivist view of cognition. Despite having gained increasing acceptance in cognitive science and cognitive neuroscience, only a surprisingly small number of cognitive architectures are fully grounded, in the sense of using only modal representations (see [64] and Section XI).

**Asynchronous operation.** Second, with the exception of two serial decision points controlled by triggers (to be described below), the *model operates entirely asynchronously*. That is, no process waits for its turn to proceed, but rather each process operates whenever its conditions are satisfied. $LIDA_I$'s asynchronous processes implement concurrent and distributed information processing, features which are vitally important when dealing with highly complex perceptual inputs with limited resources [65].

**Functional consciousness.** Third, the model includes an *explicit functional consciousness mechanism* that plays a major role in perceptual filtering, action selection, and learning (the LIDA model makes no commitment on the subject of phenomenal consciousness on the part of either animals or artificial agents [66]). Functional consciousness plays an important role as a perceptual filter by enabling the agent to focus on only the most salient information. It helps action selection by allowing the agent to recruit resources in order to choose what to do next and to efficiently solve problems. The usefulness of consciousness as viewed by GWT in enabling multiple networks in the brain to cooperate and compete has been previously argued for - see [41, 67] for examples. Some experimental data we computationally model, such as the Allport experiment [68], are difficult to account for without a functional consciousness mechanism. Moreover, the model assumes that functional consciousness is a necessary condition for learning (see Profligacy in learning below).

**Cognitive cycles.** Fourth, the *Cognitive Cycle Hypothesis*, that emerges from the LIDA model claims that human cognition functions by means of continual iteration of similar flexible cognitive cycles each taking approximately 200-300 ms [69] (similar concepts have been proposed in neuroscience [7, 8] and cognitive science [70]). However, because of cascading, cycles potentially occur at a rate of five to ten cycles per second. These cycles can cascade; that is, several cycles may have different processes running simultaneously. This cascading must, however, respect the serial nature of conscious processing that is necessary to maintain the stable, coherent image of the world that consciousness provides [71, 72]. Higher-level cognitive processes operate across multiple cycles. This view is consistent with emerging evidence from cognitive neuroscience [73, 74]. Building higher-level cognitive processes from cognitive cycles acting as "cognitive atoms" should prove a useful strategy for developing cognitive software agents, because of the computational efficiency of asynchronous and partially overlapping processes when dealing with complex information [65]. See Section X for the neuroscientific plausibility of this concept and its usefulness in accounting for experiments dealing with subjective reportability.

**Profligacy in learning**. Fifth, each of the various modes of learning in the model follows the *principle of profligacy*. This means that new representations are added to the various memories at the slightest justification, that is, whenever they come to consciousness, and are left to survive by reinforcement or they simply decay away. Such a principle is often referred to as *generate and test* because multiple representations are generated but very few survive [75]. While many cognitive architectures (including ACT-R) follow a roughly similar principle, LIDA has specific descriptions of memory systems that provide wide-ranging conceptual explanations. Examples are the more fine-grained subdivision of memory systems (Fig. 1), and the cognitively plausible auto-associative implementation of memory which can account for effects such as the tip-of-the-tongue effect or the remember-know distinction. Most of these memory systems have been implemented computationally (see Section XI), but learning is conceptual and not yet part of $LIDA_I$, with the exception of procedural learning, which has been implemented computationally in LIDA's predecessor [76], and perceptual associative learning, which is currently being implemented in a LIDA-based infant vervet monkey agent that learns predator alarm calls.

**Feelings and emotions.** Finally, the model does not have any built-in drives or specific motivators. Instead, artificial *feelings and emotions implement the motivation* needed for an agent to select appropriate actions with which to act on its environment. They also serve as primary learning facilitators by regulating the amount of reinforcement assigned to any entity in the system (see Section VII).

## III. The LIDA Cognitive Cycle

The LIDA model and its ensuing architecture are grounded in the LIDA cognitive cycle. The agent's "life" can be viewed as consisting of a continual sequence of these cognitive cycles. Each cycle consists of three phases, an understanding phase, an attending phase, and an action selection phase. LIDA's cognitive cycles closely resemble action-perception cycles in neuroscience [8, 77], and also bear some similarity to execution cycles in other cognitive architectures [10, 16, 70]. A cognitive cycle can be thought of as a cognitive "moment". As will be described in Section V below, higher-level cognitive processes are composed of many of these cognitive cycles, each a cognitive "atom."

Just as atoms have inner structure, the LIDA model hypothesizes a rich inner structure for its cognitive cycles [31, 78]. During each cognitive cycle, the LIDA agent first makes sense of its current situation as best as it can by updating its representations of both external (coming through the senses) and internally generated features of its world. This is the *understanding* phase of the cycle. By a competitive process to be described below, it then decides what portion of the represented situation is most in need of attention. This portion is broadcast to the rest of the system, making it the current contents of consciousness, and enabling the agent to choose an appropriate action to execute. This is the *attending* phase.

These broadcast conscious contents facilitate the recruiting of internal resources, potential actions, from which the action selection mechanism chooses. This is the *action* phase. Figure 1 shows the process in more detail. It starts in the upper-left corner and proceeds roughly clockwise.

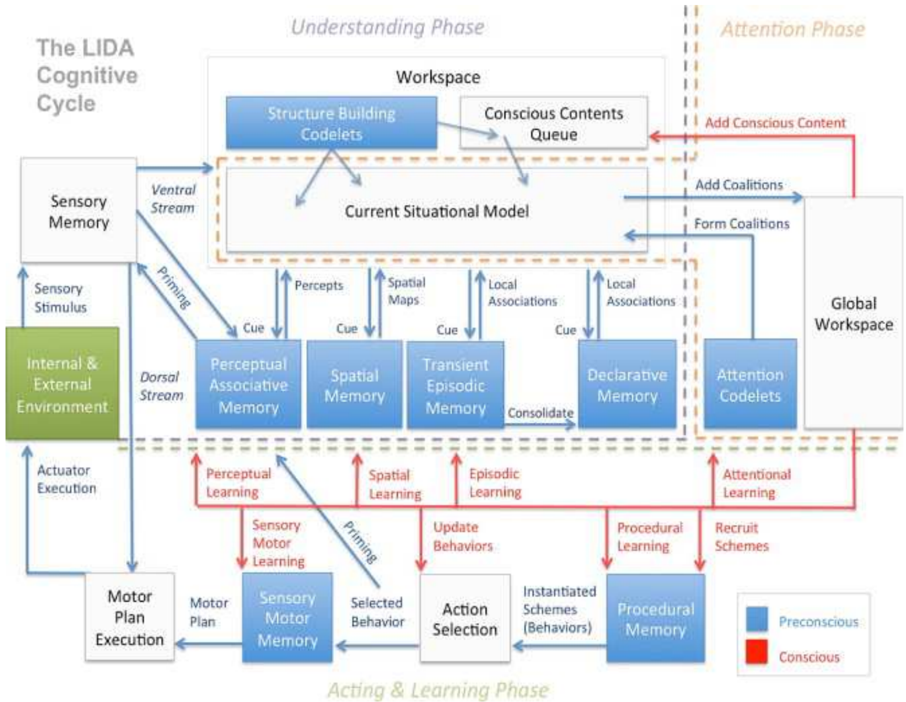


Figure 1 **LIDA Cognitive Cycle Diagram**

**Understanding phase.** The cycle begins with sensory stimuli from sources in the agent's external and internal environment being intercepted in sensory memory (e.g., the iconic buffer). Low-level feature detectors in sensory memory begin the process of making sense of the incoming stimuli. Recognized low-level features pass activation to higher-level features, such as objects, categories, relations, events, situations, etc., represented as nodes in the Perceptual Associative Memory (PAM). PAM nodes are connected by links, which can represent, for instance, correspondence, membership, spatial or causal relationships, as well as affordances, in the case of an object-action link or a category-action link [79]. These PAM nodes, and the links between them, are the building blocks of node structures in the Workspace that are similar to Barsalou's [42, 44] perceptual symbols[6] and serve as modal representations in the model (the LIDA model does not contain amodal representations). These entities, recognized preconsciously and represented by PAM node structures, make up the percept that is passed asynchronously to the Workspace, where a model of the agent's current situation, called the Current Situational Model (CSM), is assembled (updated). This percept serves as a cue to two forms of episodic memory (the memory for events), transient [80] and declarative (autobiographical and semantic). LIDA$_I$ uses auto-associative, content-addressable Sparse Distributed Memory [52] implementations to store and cue episodic, declarative and spatial memories. Responses to the cue consist of local associations, that is, remembered events from these two memory systems that were associated with the various elements of the cue. In addition to the current percept and the CSM, the Workspace contains recent percepts and the models assembled from them that have not yet decayed away. The Workspace also contains the Conscious Contents Queue, a list containing a series of a few tens of very recent conscious contents, which helps the agent to deal with time-related concepts [81]. A new model of the agent's current situation is assembled from the percepts, the associations, and the undecayed parts of the previous model. This assembling process will typically require structure-building codelets[7]. These structure-building codelets are small, special-purpose processors, each of which has some particular type of structure it is designed to build. These codelets are continually monitoring the Workspace for opportunities to fulfill their particularly specified task. They may draw upon perceptual memory and even sensory memory to enable the recognition of relations and situations, and of analogies and similarities (inspired by [50, 51]). The newly assembled model constitutes the agent's understanding of its current situation within its world. It has made sense of the incoming stimuli and the understanding phase is complete.

**Attending phase.** For an agent operating within a complex environment, this current model may well be much too rich for the agent to consider all at once in deciding what to do next. It needs to selectively attend to a portion of the model. Portions

[6]Barsalou has confirmed that our PAM node implementations are similar to his perceptual symbols in personal correspondence.

[7] In the computational model, the term codelet refers generally to any small, special-purpose processor or running piece of software code. Codelets correspond to processors in Global Workspace Theory.

of the model compete for attention. These competing portions take the form of coalitions of structures from the model. Such coalitions are formed by special-purpose attention codelets, whose function is to bring certain perceptual structures of concern to the particular attention codelet into the Global Workspace (hence the name Global Workspace Theory). The coalition containing the most salient (important, urgent, insistent, novel, threatening, promising, arousing, unexpected) perceptual structures wins the competition. In effect, the agent has decided on what to attend. A representation of the contents of the winning coalition is then broadcast globally bringing its contents to consciousness and, thereby, completing the attending phase of the cycle.

**Action and learning phase.** One major purpose of all this processing is to help the agent choose what to do next, the other being the several forms of learning. Though the contents of this conscious broadcast are available globally (facilitating different modes of learning in the conceptual model – see Section VI), the primary recipient is Procedural Memory, which stores templates ("schemes[8]") of possible actions including their contexts and possible results. It also stores an activation value for each such template which attempts to measure the likelihood of an action taken within its context producing the expected result. Templates whose contexts intersect sufficiently with the contents of the conscious broadcast instantiate copies of themselves with their variables specified to the current situation. Instantiated templates remaining from previous cycles may also continue to be available. These instantiations are passed to the action selection mechanism, which chooses a single action from one of these instantiations. The chosen action then goes to sensory-motor memory, where it is executed by an appropriate algorithm (motor plan). The action taken affects the environment, or an internal representation, or both, and the cycle is complete.

Concurrently with action selection and execution, the contents of the conscious broadcast is used to update each of several memories (Perceptual Associative (recognition), Transient Episodic, Attentional, Procedural), both by adding new items and by reinforcing existing items.

## IV. Types Of Decision Making and Action Selection

The previous sections focused on one form of action selection. Here, we discuss alternate variants, many of which have been implemented in the computational architecture.

Volitional decision making (volition for short) is a higher-level cognitive process for conscious action selection. To understand volition it must be carefully distinguished from 1) consciously mediated action selection, 2) automatized action selection, 3) alarms, and 4) the execution of actions. In each of the latter three, the actual selection (or execution) is performed unconsciously. Consciously planning a driving route from a current location to the airport is an example of deliberative, *volitional decision making*. Choosing to turn left at an appropriate intersection along a familiar route requires information about the identity of the cross street acquired consciously, but the choice itself is most likely made unconsciously - the choice was *consciously mediated* even though it was unconsciously made. While driving along a straight road with little traffic, the necessary slight adjustments to the steering wheel are typically *automatized actions* selected completely unconsciously, one action called by the previous [82]. They are usually not even consciously mediated, though unconscious sensory input is used in their execution. If a car cuts in front of the driver, often he or she will have turned the steering wheel and pressed the brake simultaneously with becoming conscious of the danger. An *alarm mechanism* has unconsciously selected appropriate actions in response to the challenge [49]. The actual turning of the steering wheel, how fast, how far, the execution of the action, is also performed unconsciously though with very frequent sensory input.

Though heavily influenced by the conscious broadcast (the contents of consciousness), action selection during a single cognitive cycle in the LIDA model is not performed consciously. A cognitive cycle is a mostly unconscious process. When speaking, for example, a person usually does not consciously think in advance about the structure and wording of the next phrase, and is occasionally even surprised at what comes out. When approaching the intersection in the example above, no conscious thought need be given to the choice to turn left. Consciousness serves to provide information on which such action selection is based, but the selection itself is done unconsciously after the conscious broadcast [36]. We refer to this very typical single-cycle process as *consciously mediated action selection*.

LIDA's predecessor IDA had computational implementations for all of the described decision-making types [30, 67]. In $LIDA_I$, computational development is still underway. Consciously mediated action selection and action execution are currently implemented as discussed below. Section V describes $LIDA_C$'s conceptual designs for the other types.

## V. Higher-level Cognitive Processes and Levels of Control in $LIDA_C$

As mentioned before, LIDA aims to be a conceptual as well as computational cognitive architecture. However, not all parts of the conceptual model have yet been implemented computationally. Sections V, VI and VII describe important parts of the conceptual model that have not yet been fully implemented. These include higher-level processes (this section), learning (section VI), and feelings and emotions (section VII).

Higher-level cognitive processing in humans includes deliberation, volition, metacognition, reasoning, planning, problem solving, language comprehension, and language production. In the $LIDA_C$ model such higher-level processes are distinguished by requiring multiple cognitive cycles for their accomplishment. They can be implemented by one or

[8] $LIDA_I$'s Procedural Memory is based on Drescher's [48] Scheme Net.

more behavior streams[9], that is, streams of instantiated schemes and links from Procedural Memory. Recall that actions (as we use the term) in the LIDA model, and in humans, include internal actions such as those used in reasoning and other higher-level cognitive processing, acting on internal representations instead of the external environment. See [34, 83, 84] for descriptions of how high-level decision making and problem solving have been implemented in $LIDA_I$. Here we focus on deliberative volitional decision making as one higher-level cognitive process.

**Deliberative Volitional Decision Making** Section IV described different forms of decision making. We now return to a consideration of deliberative, volitional decision making. In 1890, William James introduced his ideomotor theory of volition [For a recent review see 85, 86]. James postulated proposers, objectors, and supporters as actors in the drama of acting volitionally. He might have suggested the following scenario in the context of dealing with a feeling of thirst. The idea of drinking orange juice "pops into mind," that is, it is propelled to consciousness by a proposer; motivated by a feeling of thirst and a liking for orange juice, the idea becomes the contents of consciousness. "No, it's too sweet," asserts an objector. "How about a beer?" says a different proposer. "Too early in the day," says another objector. "Orange juice is more nutritious," says a supporter. With no further objections, drinking orange juice is volitionally selected.

Baars incorporated ideomotor theory directly into his GWT [22]. The $LIDA_C$ model fleshes out volitional decision making via ideomotor theory within GWT [30] as follows. An idea "popping into mind" in the $LIDA_C$ model is accomplished by the idea being part of the conscious broadcast during a cognitive cycle, that is, part of the contents of consciousness for that cognitive moment. These contents are the information (structures) contained within the winning coalition for that cycle. This winning coalition was gathered by some attention codelet (see Section III above). Ultimately, this attention codelet, by forming a coalition that wins the contest, is responsible for the idea "popping into mind." Thus we implemented the characters in James' scenario as attention codelets, with some acting as proposers, others as objectors, and others as supporters, the content of each "popping into mind" if it wins the competition and is broadcast.

But how does the conscious thought of "Let's drink orange juice," lead to a let's-drink-orange-juice node in the Workspace? Like every higher-order cognitive process in the $LIDA_C$ model, volition occurs over multiple cycles, and is implemented by a behavior stream in the action selection module. This volitional behavior stream is an instantiation of a volitional scheme in Procedural Memory. Whenever a proposal node in its context is activated by a proposal in the conscious broadcast, this volitional scheme instantiates itself. The instantiated volitional scheme, the volitional behavior stream, is incorporated into the action selection mechanism, the behavior net. The first (internal) behavior in this volitional behavior stream sets up the deliberative process of volitional decision making as specified by ideomotor theory, including writing the let's-drink-orange-juice node to the Workspace[10]. Note that a single proposal with no objection can be quickly accepted and acted upon.

This volitional decision-making process might oscillate with continuing cycles of proposing and objecting as in Eric Berne's "what if" game [87]. To counter such endless oscillations, the $LIDA_C$ model proposes three hypothetical mechanisms: reducing the activation of proposer codelets each time they reach consciousness, reducing the time allocated for the process by a "timekeeper codelet" at each restart, and a metacognitive process monitoring the process and choosing an alternative if it has gone on for too long (see [30] for details).

In addition to volition, deliberative processing is also involved in other higher-level cognitive processes such as planning, scheduling, and problem solving. Deliberative information processing and decision making allows an agent to function flexibly within a complicated niche in a complex, dynamic environment. Such deliberative processes in humans, and in some other animals, are typically performed in an internally constructed virtual reality. An internal virtual reality for deliberation requires a short-term memory in which temporary structures can be constructed with which to "mentally" try out possible actions without actually executing them. In the $LIDA_C$ model the virtual window of the perceptual scene in the Workspace serves just such a function [79]. In many cases, the action selected during almost all cognitive cycles consists of building or adding to some representational structures in the Workspace during the process of some sort of deliberation[11].

## VI. Learning in $LIDA_C$

The conscious broadcast has two primary roles: recruitment of resources, and learning. Global Workspace Theory's multi-modal learning requires that the broadcast be global, making learning critical to any understanding of GWT, which LIDA models. Learning is also critical to understanding the role played by feelings and emotions in the $LIDA_c$ model (see section VII).

Learning in the LIDA model can only occur after information has been attended to, that is, broadcast from the Global Workspace. The $LIDA_C$ model realizes several fundamental learning mechanisms (modes), each in two types, which underlie much of human learning. The two types are *instructionalist* (i.e., learning by the creation of new representations) and *selectionist* (i.e., learning by the reinforcement of existing representations) [25]. The modes of learning in the model include *perceptual*, *episodic*, and *procedural*.

[9] A stream is a sequence with its order only partially specified.

[10] Alternatively, this node could arrive in the Workspace with the percept of the following cycle as a result of internal sensing of the internal speech. In LIDA, this is only an implementation matter, making no functional difference. In humans this is an empirical matter to be decided by experiment. Thus the design decision for LIDA becomes a cognitive hypothesis.

[11] Internal actions are part of the $LIDA_C$ model, but have not been implemented yet.

Perceptual learning is learning to recognize objects, categorizations, relationships, events, etc. As new objects, categories, and the relationships among them and between them and other elements of the agent's ontology are learned, nodes (objects and categories) and links (relationships) are added to PAM, but not before the conscious broadcast (Figure 1). Episodic learning is the encoding of information into episodic memory, the associative, content-addressable, memory for events - the what, the where, and the when [88, 89].

Procedural learning is the encoding of procedures for executing behaviors into Procedural Memory. It is the learning of new actions and action sequences with which to accomplish new tasks. Here we must distinguish between action selection and action execution. LIDA's Procedural Memory is composed of schemes concerned with the selection of actions. Algorithms (motor plans) for their execution are found in Sensory-Motor Memory. The Procedural Memory has been implemented in $\text{LIDA}_{\text{I}}$, and a procedural learning implementation was available in IDA [76].

Instructionalist learning refers to learning by the addition of new representations. For perceptual learning, such new representations are produced in the Workspace by structure-building codelets. If a new representation is part of a winning coalition in the Global Workspace, it becomes part of the conscious broadcast and is learned. In the current implementation, for perceptual learning these new representations will consist of nodes and links in PAM, for procedural learning of schemes in Procedural Memory, and for episodic learning of vectors in Transient Episodic Memory.

Each node in PAM and each scheme in Procedural Memory has both a base-level and a current activation. The current activation measures the present relevance or saliency of the node or scheme. Their base-level activation measures their overall usefulness. Occurring during each conscious broadcast, selectionist learning reinforces the base-level activation of every node and scheme in the conscious content of the broadcast. For episodic learning, such reinforcement happens automatically by means of internal mechanisms of sparse distributed memory (SDM) [52], the computational mechanism we use to model episodic memory.

Although the types of knowledge retained due to these three learning mechanisms differ, we hypothesize that conscious awareness is sufficient for learning. Although subliminal acquisition of information appears to occur, the effect sizes are quite small compared to conscious learning. In a classic study, Standing [90] showed that 10,000 distinct pictures could be learned with 96% recognition accuracy, after only 5 seconds of attention to each picture. No intention to learn was needed. Consciously learned educational material has been recalled after 50 years [91]. Attention greatly facilitates most modes of learning.

All learning in LIDA occurs as a result of the conscious broadcast. The conscious broadcast contains the entire content of consciousness including the affective portions. Transient Episodic Memory is also updated with the current contents of consciousness, including *feelings*, as events (episodic learning). *Up to a point, the stronger the affect is, the stronger the encoding in memory* (discussed in more detail in the next section). Procedural memory (recent actions) is updated (reinforced) with *the strength of the reinforcement influenced by the strength of the affect* (procedural learning).

Most of LIDA's learning mechanisms are conceptual at this stage. Implementations exist for procedural learning [76] and episodic learning [92]. Spatial learning is currently being developed for possible robotic applications of LIDA. Perceptual associative learning is currently being implemented in a LIDA-based infant vervet monkey agent that learns predator alarm calls. Additional modes of learning are in the planning stage for later implementation into the $\text{LIDA}_{\text{I}}$ architecture. These include learning of motor plans in Sensory-Motor Memory for the execution of actions, the attentional learning of new attention codelets, and the learning of new structure-building codelets.

## VII. Feelings and Emotions in $\text{LIDA}_{\text{C}}$

Emotions have been argued to play major roles in facilitating high-level cognition (for example, by acting as motivators for actions): "the emotional aspect of cognition, providing motivation and value to an otherwise neutral world, [...] is a fundamental part of the make-up of an organism with respect to sensorimotor learning"[93]. However, the modeling of emotion has been largely neglected in cognitive architecture research [4], with notable exceptions including SOAR [94, 95], and models accounting for emotion as well as some other aspects of cognition, but not aiming to be comprehensive architectures (e.g. [96], see [97, 98] for reviews). In this section, we will describe how $\text{LIDA}_{\text{C}}$ can model emotions and use them as motivators for action selection. These ideas are part of conceptual $\text{LIDA}_{\text{C}}$ and have not yet been implemented in $\text{LIDA}_{\text{I}}$. As with the other cognitive science phenomena described in this paper, their implementations will be simplified abstractions intended to be functionally similar to the real phenomena. We do not claim to account for their full psychological complexity. Nevertheless, using terms common in cognitive science is useful for establishing conceptual grounding, and to reduce the need for explanations.

The word "feeling" may be associated with external haptic sense, such as the feeling in our fingertips as they touch the keys. It is also used in connection with internal senses, such as the feeling of thirst or the pain of a pinprick. Following Johnston [99], and consistent with the influential appraisal theory [100], in the $\text{LIDA}_{\text{C}}$ model we speak of *emotions* as feelings with cognitive content, such as the joy at the unexpected meeting with a friend or the embarrassment at having said the wrong thing.

Contemporary theories of emotion posit that cognitive *appraisals* of physiological changes give rise to emotional states [101-104]. Appraisal is an unconscious or conscious process where emotions are produced from subjective evaluations of situations, or objects, or events, along dimensions such as novelty, goal-alignment, agency, coping potential, and availability of a plan.

**Representing feelings in $\text{LIDA}_{\text{C}}$.** Feelings are represented in the LIDA$_C$ model as nodes in PAM. Each feeling node constitutes its own identity, for example, distress at not having enough oxygen is represented by one node, relief at taking a breath by another. Each feeling node has its own valence, always positive or always negative, with varying degrees of arousal. The current activation of the node measures the momentary arousal of the valence, that is, how positive or how negative. The arousal of feelings can be bottom-up, that is, arising from feature detectors in Perceptual Associative Memory (PAM). If those feeling are also emotions, the arousal can also be top-down, that is influenced by the appraisal that gave rise to the emotion. A thirst node in humans would activate itself in response to internal sensation having to do with fluid balance. A fear node in the presence of a known event would be activated by spreading activation from the other nodes representing the event, in turn activated by feature detectors of different sensory modalities.

Like other Workspace structures, feeling nodes help to cue transient and declarative episodic memories. The resulting local associations may also contain feeling nodes associated with memories of past events, which is consistent with network theories of emotion [105]. Being part of the structure carried by the coalition, and bringing their own activation with them, these feeling nodes play a major role in assigning activation to coalitions of information to which they belong, helping them to compete for attention. Any feeling nodes that belong to the winning coalition become part of the conscious broadcast (i.e., part of the contents of consciousness, and can influence the selection of an appropriate action).

Feelings can be recognized based on sensory input. Taking thirst as an example as we did above, an internal sense may sufficiently activate the thirst node in PAM, causing an instantiated thirst node to appear in the LIDA$_C$'s Workspace. If this node is selected by an attention codelet and the resulting coalition wins the competition in the Global Workspace and thus comes to consciousness, the feeling of being thirsty is experienced[12] by the agent.

As stated above, emotions in LIDA$_C$ are taken to be feelings with cognitive content [99], for example, the fear of a truck bearing down, the shame at something said, the sadness at a loss, or the surprise at an unexpected turn of events. Feelings, including emotions, are represented by nodes in LIDA$_C$'s Perceptual Associative Memory (PAM). Cognitive content, represented by node/link structures, are linked to emotion nodes by a process called appraisal [106].

**Appraisal in LIDA.** In LIDA$_C$, appraisal of a new event, and its connection to an appropriate emotion, is performed by appraisal codelets[13], a form of structure-building codelet, acting within LIDA$_C$'s Workspace. Appraisal codelets identify an emotion as well as an arousal level (see below), in the form of an emotion PAM node, and connect this node to the perceptual structure representing the event causing the emotion. The appraisal process can also alter previously identified emotions when the event is reappraised. This newly appraised structure, including the emotion node, is incorporated into the Current Situation Model (CSM), from whence some attention codelet may succeed in bringing it to consciousness. The conscious emotion can subsequently motivate action selection.

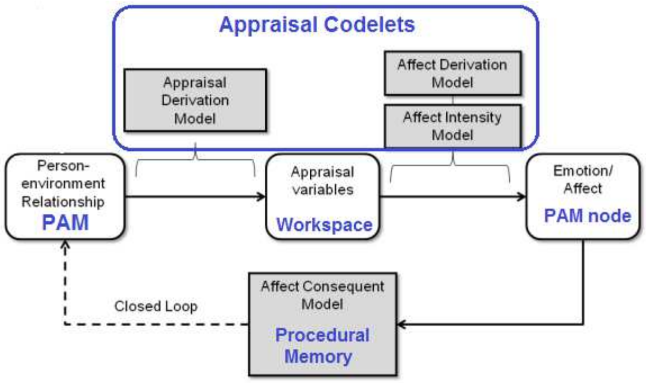


**Figure 2. Components of appraisal models (black) - based on [98] – and how LIDA$_C$ accounts for them (blue)**

LIDA$_C$'s appraisal process is based on components of computational appraisal models proposed by Marsella et al. [98] (see Figure 2). Here, person-environment relationship refers to the representation of the agent's current relationship with its environment. In LIDA$_C$'s case, this representation is built in the Workspace, taking the form of PAM nodes.

Appraisal variables are derived from this representation and mapped onto an affective state (an emotion) with a specific intensity by an *affect derivation* model and an *affect intensity* model. In LIDA$_C$, this is done by appraisal codelets operating on the Workspace. We propose to use Scherer's [100] appraisal variables: relevance, implications, coping potential and normative significance (all subjective to the current person-environment relationship). In LIDA$_C$, these variables are represented in the Workspace by PAM nodes, node structures, and their activations (see below). An emotion PAM node is spawned and connected to this representation by an appraisal codelet. Thus, these codelets implement the Appraisal Derivation Model and the Affect Derivation Model in Marsella et al.'s terminology [98] by creating the emotion node, and the Affect Intensity Model by creating and weighting the links to the Workspace node structures representing appraisal variables, which will ultimately determine the activation (or arousal) of the emotion node.

According to Scherer [100], the *relevance* of a stimulus event to an organism can be judged by its novelty (which includes intensity, suddenness, familiarity and predictability), as well as intrinsic pleasantness and goal relevance. In LIDA$_C$, intensity or bottom-up salience is implemented by feature detectors, intrinsic pleasantness by activation passed from pleasant or unpleasant emotion nodes, and top-down importance with regard to current goals by activation passed down from goal representations.

The *implications* of a situation or an event need to be appraised to determine whether it furthers or hinders an

[12]In the sense of functional consciousness. We make no commitment in our LIDA model to phenomenal consciousness.

[13] This section describes a purely conceptual model. Appraisal codelets and variables have not yet been implemented in LIDA$_I$.

organism's survival, or its ability to satisfy its needs or attain its goals. This involves the attribution of a cause, the assessment of outcome probabilities and of discrepancies from expectations, and checking the conduciveness to goals or needs. In $LIDA_C$, implications could be judged by a predictive mental model in the virtual window of the CSM in the Workspace [79], represented as a node structure, that might be created and altered by behavior schemes representing the dynamics of the cause (e.g. possible actions of the responsible person) in the Procedural Memory [107]. This model or node structure would consist of PAM nodes (which can represent events as well as entities or objects [79]) representing causes and outcomes, and their activations would represent outcome probabilities and utilities. The overall activation of the node structure would influence the activation of the emotion node (urgent implications would lead to an emotion node with high activation).

*Coping potential* involves the evaluation of whether the individual can affect its concern with the eliciting event, and depends on to what extent the event can be controlled or influenced by the agent as well as to what extent the agent can adjust or adapt to the consequences of the event. In $LIDA_C$, this also could be evaluated using a model created in the virtual window of the CSM [79], similarly to the implication evaluation, in this case making use of learned schemes representing the agent's own actions. This evaluation might require multiple cognitive cycles, each selecting a possible action, adding its results to the model in the Workspace and evaluating whether and to what extent the eliciting event will have been dealt with; this extent will then influence the activation of the emotion node. In both implication evaluations and coping potential evaluations, if a similar event has been encountered and its consequences learned already, building a predictive model using such schemes might not be necessary – the event's consequences can be cued from episodic (or perceptual) long-term memory.

Finally, *normative significance* involves taking into account how other group members would react to an event, and the significance of this for social agents [100]. The normative significance of events is evaluated against internal as well as external (social) standards. Such standards could be represented in semantic memory (part of declarative memory) in $LIDA_C$, and, if cued, could influence the appraisal of socially significant situations, either by modulating the activations of the node structures representing these or by adding additional nodes (see previous work on moral standards in LIDA in [108]).

Importantly, we hypothesize that none of these appraisal variables require any amodal representations, as is common in other computational models of emotion (see e.g. [98]). All of them are represented by PAM nodes (which are based on perceptual symbols [42, 44]) and their activations.

Based on these appraisal variables, appraisal codelets can assign an emotion to the appraised situation, i.e. they can build a node structure representing the situation as well as its appraisal (both of which consist of PAM nodes) and connect an emotion PAM node to this structure. The activation of the emotion node (i.e. the intensity of the represented emotion) will be derived from this node structure, and will depend on all the factors described above.

The affect consequence model, mapping the affect onto an either behavioral or cognitive (internal) change, is implemented by the Procedural Memory and Action Selection modules in LIDA, which can cause the selection and execution of an external (behavioral) or an internal action. These actions cause changes in the represented situation in the Workspace, which is used in subsequent appraisals. Thus $LIDA_C$ contains a closed-loop appraisal system.

Although $LIDA_C$ adopts an appraisal-model of emotion, it has two major differences in comparison to recent computational models of emotion reviewed by Marsella [98]. First, our model may potentially account for more factors determining the intensity of emotions than conventionally used affect intensity models such as e.g. expected utility (intensity proportional to the product of goal utility and probability of attainment – see [98]), since the node structure resulting from the appraisal process and passing activation to the emotion node could possibly be highly complex. This method of deriving affective intensity is also arguably more cognitively plausible than using a mathematical equation and amodal symbols (e.g. [42, 44]). However, since our model is purely conceptual as of yet, these claims are speculative and require further computational testing. Second, $LIDA_C$'s attention mechanism provides computational explanations for the demonstrated importance of attention in the subjective intensity of emotions (e.g. [109, 110] – an agent paying attention to an emotion has an attention codelet with a high activation that will build a coalition with said emotion, increasing its activation and thus its subjective intensity). The few emotion models accounting for attention (e.g. [111]) usually only include a basic thresholding mechanism, as opposed to $LIDA_C$'s detailed attention model that is based on Global Workspace Theory [69, 112].

**The role of emotions in action selection**. Every autonomous agent must be equipped with primitive motivations that motivate its selection of actions, in order to form its own agenda [6]. Such motivations may be causal as in the purely physical mechanism motivating a bacterium to follow a nutrient gradient upstream [113]. They may occur as drives as in the *if* condition of a production rule in an artificial agent [114]. In humans, in many animals, and in the $LIDA_C$ model, these motivations are implemented by feelings and emotion [33]. Such feelings implicitly give rise to values, an agent's general preference for an action in a situation, that serve to motivate action selection [33, 115]. Feelings provide flexible motivations. For example, *hunger* with its multiple, learned satisfiers is much more flexible than specifying under which circumstances to eat what. Also, a built-in or learned fear of A can be flexibly applied to B when B is like A. Feelings are desirable motivators when the environment is too complex to specify what to do when, and when association and learning are both available.

LIDA's Procedural Memory contains schemes, each consisting of a context, an action, a result, and an activation measuring the likelihood of an action taken in the context producing the result. Feeling or emotion nodes in the conscious broadcast that also occur in the context of a scheme in Procedural Memory add to the current activation of that scheme, increasing the likelihood of it being activated. It is here that feelings play their first role as implementations of motivation by adding to the likelihood of a particular action being selected. A feeling in the context of a scheme implicitly increases or decreases the value assigned to taking that scheme's action.

Apart from facilitating action selection, feelings or emotions in the conscious broadcast also play a role in modulating the various forms of learning. Up to a point, the higher the arousal the greater the learning. Beyond that point, more arousal begins to interfere with learning [116].

In the Action Selection mechanism, the activation of a particular behavior scheme, and thus its ability to compete for selection and execution, depends upon several factors. These factors include how well the context specified by the behavior scheme agrees with the current and very recently past contents of consciousness (that is, with the contextualized current situation). The contribution of feeling nodes to the behavior scheme's activation constitutes the environmental influence on action selection. As mentioned earlier, the activation of this newly arriving behavior also depends on the presence of feeling nodes in its context and their activation as part of the conscious broadcasts. Thus feelings contribute motivation for taking action by adding activation to newly arriving behavior schemes.

The selected behavior, including its feelings, is then passed to sensory-motor memory for execution. There the feelings modulate the execution of the action [117].

## VIII. LIDA AND THE UNDERLYING NEURAL PROCESSES

As emphasized earlier, LIDA is *not* intended to model the neural mechanisms underlying cognition. But if LIDA is to be a cognitive model, and cognition is implemented in brains, there must be some relationship between LIDA and the underlying neuroscience. In this section, we will outline this relationship, in order to argue for the plausibility of the LIDA model, and to further clarify and constrain the functionality that LIDA's modules and processes are intended to model. Following Freeman and others we invoke non-linear dynamics as the needed bridge between our model and the underlying neuroscience [118-120]. Although currently not implemented as a dynamical system, LIDA's cognitive cycle shows many of the properties of such systems [121]. For example, it is similar to an overarching oscillatory process, and is assembled from multiple components themselves resembling oscillators; its dynamics change over multiple time scales (from activation-passing processes operating in a few ms, to modules operating in a few tens of ms, to cognitive cycles and multi-cycle processes [69]); and its representations show properties resembling the dynamic systems concept of stability (a PAM node representation in the Workspace might be stable, i.e. persistent in the face of systematic or random perturbations - incoming activations, or unstable if it is still in PAM and its activation is very close to the percept threshold) [122].

Descriptions of tentative neuronal correlates for LIDA's modules and processes, based on functional correspondence, have been described elsewhere [123]. Below, we shall briefly outline a simplified dynamical systems view of LIDA's cognitive cycle, connecting it to empirical neuroscience. We use a flavor of dynamic systems theory called Dynamic Field Theory (DFT) to make this connection, because of its conceptual similarity to LIDA's ideas (see above) and its neurobiological plausibility [122, 124, 125]. Dynamic neural fields in DFT can be viewed as types of recurrent neural networks [126] with dynamics similar to leaky integrate-and-fire equations which are used in some spiking neuron models [127]. They are firmly grounded in neural principles [126] and can account for the dynamics of cortical neural activation, for example in the visual cortex [128] and motor cortices, substantiated by comparisons of single-neuron recordings to the field activation [129] (see [126] for more arguments for the neural plausibility of DFT).

Dynamic neural fields formalize how neural populations represent dimensions characterizing perceptual features, cognitive operations, actions, etc. They represent information along an activation dimension (corresponding to the amount of available information or its certainty) and one or more feature dimensions (e.g. spatial location, orientation, or perceptual features such as frequency or color or motion), with low levels of activation at a particular point indicating that that value of the represented dimension is unlikely, and with the dynamics defined by a field equation similar to the one below [122].

$$\tau\dot{u}(x,t) = -u(x,t) + \text{resting level} + \text{input} + \text{interaction}$$

where $u(x,t)$ is the activation field defined over dimension x and time t, and $\tau$ is a timescale parameter. Without the interaction between field sites (ignoring the last term), attractor solutions depend only on the field input (e.g. from sensors) and the constant resting level; activation peaks would eventually vanish with ceasing input. To stabilize local activation peaks in the absence of input, the interaction is defined to be locally excitatory and globally inhibitory [122], in the center-surround fashion observed in biological neurons (e.g. [130]).

Different layers of LIDA's sensory memory [131] could correspond to sensory cortical areas, which could be modeled as multiple dynamic fields – e.g. [125]. Multiple such features represented on different fields implementing different layers of sensory memory can be bound into holistic object representations on a working memory field, e.g. as done by [132]. Such a field might implement the Workspace, and activation peaks on it correspond to PAM nodes (Figure 3).

Another dynamic field with strong inhibitory interaction to ensure a winner-take-all mechanism could implement LIDA's Global Workspace. This field would receive its excitatory input from the Workspace field, as well as an Attention field for amplifying attended regions, and would select and stabilize the strongest activation peak, which would then inhibit all others and emerge as the winner of the competition for consciousness (see Section III). Adjusting the timescale parameter can ensure that this process only happens at plausible frequencies.

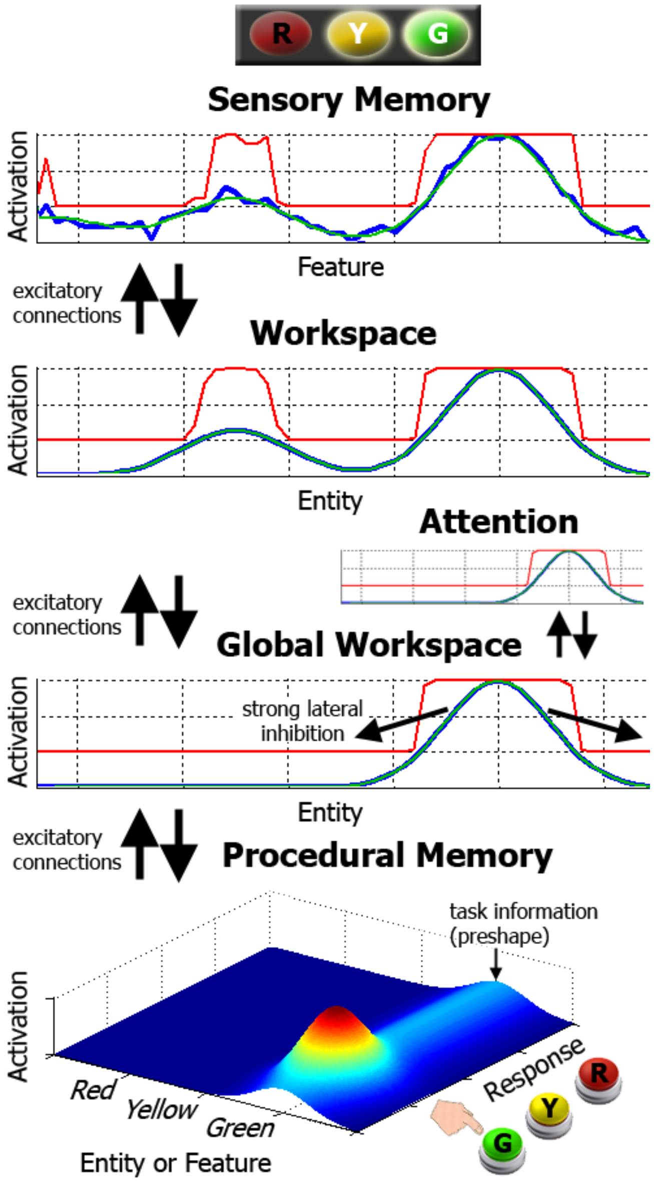


**Figure 3. A dynamical systems (DFT) based implementation of the LIDA cognitive cycle in a reaction time experiment**. The blue, thick lines / blue surface represent dynamic neural fields. Red, thin lines show the field output, after the application of a sigmoidal threshold function to ensure that only sufficiently activated field sites contribute to the interaction [122]. In this simplified example, yellow and green glowing lights are perceived and represented (only one feature layer is shown, but any number could be combined in Workspace representations). The activation peak representing the green light enters the Global Workspace, due to the Attention field amplifying its peak (but not the peak associated with the yellow light), and elevates the pre-shaped Procedural Memory field above threshold, causing an action to be taken and the button to be pressed (pre-shaping can ensure that from all applicable actions based on current percepts, only the task-relevant one reaches the threshold, and has been argued to be neurally plausible [129]).

Finally, action selection could be implemented as suggested by [133], by adding two additional layers, apart from the Global Workspace layer, representing important objects: a Procedural Memory (called ASL in [133]) encoding chains of action primitives, and a Result layer (called IL in [133]) specifying result states of an action. Conscious object representations can then prime multiple possible actions (activation peaks) in the Procedural Memory layer, and a predefined task context or a specific goal defined in the Result layer can then elevate one of these over the threshold, stabilizing it and thus selecting the action best matching the preconditions and the goal state. Action selection and movement control using dynamic neural fields [134], as well as pre-defining goals through pre-shaping [129], have been argued to be neurobiologically plausible.

Figure 3 shows the described highly simplified DFT implementation of $LIDA_I$ for the reaction time experiment described in Section X, showing the agent perceiving a green light and pressing the appropriate button. Since we have only recently started investigating a dynamic implementation of $LIDA_I$, we do not yet have data substantiating the neural plausibility of such an implementation, or of our mapping of LIDA modules to brain areas [123].

LIDA's modules can be tentatively assigned neural correlates, based on functional correspondences. This mapping is tentative because the empirical neuroscience evidence is still changing. Such correlates should be interpreted as being involved with the activity of the corresponding LIDA module or process, rather than as being equivalent to it. The correspondence of module to brain correlates is often one-to-many, since a single LIDA module may be implemented by numerous, disparate cell assemblies in brains. Do note that the LIDA model, being described in terms of modules and processes, makes no commitment to the underlying neural structure being modular or localized, as is exemplified in the following paragraphs.

Sensory memory correlates in a one-to-many fashion with brain areas for each sensory modality, for example iconic memory (occipital lobe) and echoic memory (primary auditory cortex) [135, 136]. Representations in Sensory Memory are all modal. Node representations in PAM (LIDA's perceptual symbol implementations [42, 44]) are difficult to localize in the brain, since they are distributed and multimodal [42, 137, 138]; some of the major areas involved are the perirhinal cortex [139, 140], amygdala and orbito-frontal cortex [141], mushroom body neurons [142], medial orbitofrontal cortex [143], etc.

The entorhinal cortex, together with the temporo-parietal and frontal lobes in humans, would implement parts of the LIDA Workspace (preconscious working memory buffers) where new objects, events, and higher-level structures are created [144, 145]. We view the hippocampus as implementing LIDA's Transient Episodic Memory [146, 147],where events are encoded in humans for a few hours or a day at most [148], as well as the Spatial Memory module [149, 150] which is currently in development.

LIDA's Global Workspace, where coalitions compete for attention, can be thought of as possibly corresponding to different brain areas during successive cognitive cycles, with each such area holding a coalition appropriate to it, Coalitions race to threshold with the winner "igniting" [151] to give rise to a thalamocortical core [152], which implements the conscious broadcast via a dynamic Global Workspace (dGW), presumably facilitated by large scale oscillatory synchrony [69, 153, 154]. LIDA's Procedural Memory would correspond to the anterior cingulate and the striatum [155, 156] while its Action Selection mechanism would be grounded in the basal ganglia [157]. These last two modules are concerned with what action to perform in response to the situation understood during a single cognitive cycle. The correlates of volitional decision making, arising from multiple cognitive cycles (see next section), include the ventral anterior cingulate cortex and prefrontal cortices (such as the ventromedial and dorsolateral prefrontal cortex) [158, 159]. Although these prefrontal areas are involved in many other tasks, their importance for volitional decision making is highlighted by the apparent necessity of prefrontal involvement in the oscillatory synchrony giving rise to conscious activity [153, 160], which is necessary for selecting volitional actions (see Section IV). LIDA's Sensory-motor Memory, which is concerned with how to perform the selected action, would involve the dorsal striatum [161]. For a more comprehensive overview of neural mappings of LIDA modules and processes see the tables in [123].

In contrast to LIDA, some other cognitive architectures and models have attempted to directly correlate the activity in their modules to brain areas, and have presented evidence for the neuronal counterparts of their modules based on brain imaging data [162]. For example, ACT-R has been successful in predicting fMRI activation in tasks such as algebraic problem solving [163] or mathematical reasoning [164]. The neuronal correlates underlying ACT-R's modules that have been substantiated using such fMRI studies include the fusiform gyrus (visual), posterior parietal cortex (imaginal), anterior cingulate cortex (goal/control), lateral inferior prefrontal cortex (retrieval), caudate nucleus in the basal ganglia (procedural) and the motor cortex (manual) [165, 166]. Apart from substantiation of their claims with neural imaging methods, another difference between LIDA's and ACT-R's neural mapping is that ACT-R assumes a strictly modular organization of the brain, with functional one-on-one mappings to individual areas, a view that has some challenges based on empirical results [137, 167, 168]. Anderson et al. [166] also point out that there is some evidence against the prediction arising from this mapping that the basal ganglia (as the counterpart of ACT-R's production system) is the sole path of communication between different cortical areas. This evidence includes observed cortical-to-cortical connections (e.g. [137, 169]), the apparent small-world properties of the cortex (minimal-length pathways connecting all individual components) [170, 171], as well as the major role of long-range synchronization of oscillatory signals in mediating communication between different cortical networks [154, 172]. Finally, analyzing cognitive states using stimulus-locked averaging across trials, as done in many brain imaging studies including ACT-R's, removes information about how the brain's spontaneous activity interacts with stimulus-driven input [172] by averaging out signals not time-locked to the stimulus.

In contrast, dynamical system properties outlined above, as well as oscillatory activity and brain rhythms, play a major role in LIDA's view of the neuronal correlates underlying its modules and processes [67, 69, 173, 174]. LIDA itself is modular, but does not try to map its modules to brain areas in a one-on-one fashion. Based on the cognitive cycle hypothesis (Section II) and the assumption that functional consciousness requires large-scale theta-gamma synchrony [154, 173], we have derived the temporal length of the cognitive cycle and its subprocesses [69], and have used these parameters to replicate behavior data (Section X). Efforts are under way to further substantiate the hypotheses put forward by the LIDA model with respect to its neuronal and oscillatory counterparts.

## IX. The $\text{LIDA}_\text{I}$ Computational Framework

The $\text{LIDA}_\text{I}$ Framework is a generic and customizable computational implementation of aspects of the $\text{LIDA}_\text{C}$ model, programmed in Java. The main goal is to provide a generic implementation of the model, easily customizable for specific problem domains, so as to allow for the relatively rapid development of $\text{LIDA}_\text{I}$-controlled software agents. Here, we briefly describe the $\text{LIDA}_\text{I}$ Framework elements and principal characteristics. A more detailed description can be found in [175].

The Framework permits a declarative description of the specific implementation details of an agent. The architecture of the software agent is specified using an XML-formatted file called the *agent declaration file*. In this way, the developer does not need to define the entire agent in Java; he or she can simply define it using this XML specification file. For example, new behaviors (schemes) can be added to an agent manually by entering a new entry to the parameters of the Procedural Memory module, and specifying the PAM nodes constituting the scheme context (in which situation the action would be appropriate), the action that should be taken when the context is matched, the expected result of the action, and a base-level activation (see Sections III and VI).

An important goal of the Framework is its ready customization. The customization can be done at several levels depending upon the required functionality. At the most basic level, developers can use the agent declaration file to customize their applications. Several small pieces in the Framework can also be customized by implementing particular versions of them. For example, new strategies for decaying activations or types of codelets can be implemented. Finally, more advanced users can also customize and change internal implementation of whole modules. In each case, the Framework provides default implementations that greatly simplify the customization process.

The Framework was conceived with multithreading support in mind. Biological minds operate in parallel and so should artificial ones, not only for plausibility, but also in order to be able to deal with complex perceptual information with limited resources[14] [65]. Tasks, encapsulations of small processes, together with a dedicated task manager, implement multithreading support that allows for a high level of parallelization. Finally, the $LIDA_I$ Framework implementation adheres to several important design principles [176] and best programming practices.

The $LIDA_I$ Framework defines several data structures and procedures (algorithms) and is composed of several pieces. Its main components are software modules, interconnected elements that represent conceptual modules in the $LIDA_C$ model. Each main component of the $LIDA_C$ cognitive model has a corresponding module in the framework. For example, the Sensory Memory, Workspace and Action Selection are all modules in the framework. In addition to a common application programming interface (API), each module has its own API that defines its functionality. Modules can have submodules. A submodule is a module nested inside another module. For example, the Workspace has several submodules, such as the CSM submodule.

Most modules in the $LIDA_I$ Framework are domain independent. For each of these modules, the Framework provides a default implementation. For example, Transient Episodic Memory is implemented as sparse distributed memory [52] and the Action Selection Module as a behavior net [56]. However, some modules must be domain specific. In particular, Sensory Memory and Sensory-Motor Memory have to be specified on the basis of the domain that the Framework is being applied to. Nevertheless, the Framework supplies base implementations from whence the developer can implement domain-specific functionality.

Modules need to perform several tasks in order to achieve their specific functionalities. The Framework provides Tasks, which are encapsulations of small processes. A module can create several Tasks to help it perform its functions. A Task can run one time or repeatedly. A Task that passes activation is an example of the former, while a structure-building codelet is an example of the latter. The Task Manager controls the execution of all Tasks in the Framework. Tasks can be executed on separate threads by the Task Manager, achieving parallel execution in a way that is approximately transparent to the user.

Modules need to communicate with other modules. To implement this, we use the Observer design pattern [176]. In short, a module, called the "listener," which receives information from another "producer" module, can register itself with the producer as a listener. Each time the producer has something to send, it transmits the information to all of its registered listeners. There are numerous instances of listeners in the Framework. One module can be registered as a listener of several other modules. Also a module can be a producer and a listener of other modules at the same time.

Nodes, links, and other LIDA elements such as coalitions, codelets, and behaviors, have activation. The activation generally represents the relative importance of the element to the current situation. All activations are excited or decayed using "strategies." These are implementations of the strategy design pattern which allows for customizable behavior; in this case they specify the way activation of each element is excited or decayed, so it is easy for the developer to change the algorithm for excitation or decay of elements.

Finally, the Framework includes several supporting tools, such as a customizable graphical user interface (GUI), logging capabilities, and an architecture loader that parses several XML files with the definition and parameterization of the agent.

Vector LIDA is a promising improvement of the LIDA cognitive architecture's computational implementation. Vector $LIDA_I$ [177] employs high-dimensional vectors and reduced descriptions. High-dimensional vector spaces have interesting properties that make them attractive for representations in cognitive models [178]. The distribution of the distances between vectors in these spaces, and the huge number of possible vectors, allow noise-robust representations where the distance between vectors can be used to measure the similarity (or dissimilarity) of the concepts they represent. Moreover, these high-dimensional vectors can be used to represent complex structures, where each vector denotes an element in the structure. However, a single vector can also represent one of these same complex structures in its entirety by implementing a *reduced description*, a mechanism to encode complex hierarchical structures in vectors or connectionist models [179]. These reduced description vectors can be expanded to obtain the whole structure, and can be used directly for complex calculations and procedures, such as making analogies, logical inference, or structural comparison [see 178, 180 for further discussion of these applications].

Vector LIDA will utilize a new reduced representation model, the Modular Composite Representation (MCR) based on high-dimensional integer vectors [181]. This representation has advantages over previous similar models: it has good representation capability with relatively simple operations (see [181] and [182] for details). Also, several new variations of Sparse Distributed Memory (SDM) [52], the Integer SDM [183] and the Extended SDM [184] provide support for storing these vectors with an intrinsic learning mechanism.

This new implementation will present several advantages over the current version. First, MCR vectors have the potential of directly implementing Barsalou's perceptual symbol system [44]. Constructing MCR vectors from sensory and motor information using hyperdimensional computing operations would produce representations that have many of the perceptual symbols' characteristics described by Barsalou [44]. Similar sensory information would yield similar representations, and the processing operations of MCR could facilitate the implementation of the simulators described by Barsalou, such as integrating details [185], simulating event sequences [184], and categorizing new stimuli [52]. Second, many cognitive operations require approximate comparisons, which are hard to implement with graph-like representations, but are natural for vector representations. Third, Integer SDM

[14] The implementations based on $LIDA_I$ are not yet at a stage where the functional importance of parallel and asynchronous operation could be verified. Ongoing work on implementing $LIDA_I$ on robots might make such empirical evaluations possible in the future.

and Extended SDM would provide an inherent learning mechanism [52] that will reinforce common vectors in the memory. Finally, the vector nature of this model makes it a good candidate for parallel implementations, using GPUs or other high-performance parallel architectures.

Although this new implementation is still in progress, extensive research and implementations have already been carried out for its main components: MCR vectors and the various SDM implementations.

Summing up, the $LIDA_I$ Framework allows the creation of new applications and simulations based on the $LIDA_C$ model. Its design and implementation aims at simplifying this process by permitting the developer to concentrate on the specifics of the application, while hiding the complexities of the generic parts of the model. Use of the Framework also enforces good software practices that simplify the creation of complex architectures. It achieves a high level of abstraction permitting several ways and levels of customization with a low level of coupling among modules. Supplemental tools like a GUI and logging support are also provided. The result is a powerful and customizable tool with which to develop $LIDA_I$-controlled software agents and robots.

## X. $LIDA_I$-BASED SOFTWARE AGENTS

We have developed four cognitive software agents that replicate experiment data from human subjects [24, 69] in order to show how the computational $LIDA_I$ architecture can model human cognition in basic psychological tasks. Our main goals with these agents were to substantiate some of the claims of the LIDA model and to take a first step towards identifying a set of internal parameters. Ideally, these internal parameters will remain constant when disparate datasets from different experiments conducted on human subjects are reproduced with $LIDA_I$ agents. Finding such a set of parameters would provide substantial evidence of the accuracy and usefulness of the conceptual cognitive model.

Basic values for the parameters governing mechanisms in $LIDA_I$ were derived from neuroscience data [69]. For example, visual feature detectors in $LIDA_I$ agents have to take about 30ms to run, derived from neuronal conduction delays in area V1 in the human visual cortex [186, 187]. These basic parameters were first tested in a simple reaction time task ($LIDA_I$ Reaction Time agent), and verified in an experiment designed to investigate perceptual simultaneity and continuity ($LIDA_I$ Allport agent), and two experiments examining the properties of attention (the $LIDA_I$ Attention and Attentional Blink agents). The latter three agents were also motivated by the goal of validating some of the claims of the GWT of consciousness underlying the LIDA model. GWT posits that consciousness is discrete, which is consistent with some recent neuroscientific evidence [154, 188, 189].

**The $LIDA_I$ Reaction Time agent**. The $LIDA_I$ Reaction Time (LRT) agent performs a simple reaction time task. The main goal was to evaluate the behavioral consequences of the parameters derived from empirical neuroscience evidence, concerning the duration of the cognitive cycle and its phases. The agent is embedded in a simple environment consisting of a red or green light, and a button which the agent has to press when the light turns green.

The LRT agent is based on the $LIDA_I$ computational Framework and contains additional code to implement the simple environment [69]. Some parts of the understanding phase of the LIDA cognitive cycle (Transient Episodic Memory, Declarative Memory, structure building codelets) were not required because of the simplicity of this task.

The LRT agent's cognitive cycle starts with a representation of the light in the environment, which is stored in Sensory Memory. Feature Detectors pertaining to the color of the stimulus observe this representation and pass activation to corresponding PAM nodes, which are then copied to the Workspace, indicating that the stimulus has been recognized or *understood* (this occurs in about 100ms, as in humans [190]). This marks the end of the *understanding phase,* which in more complex domains would also include memory recall and structure building[15].

In the *attending phase*, attention codelets look out for relevant, important, urgent, or novel percepts, combine them into coalitions, and move them to the Global Workspace. One of these coalitions wins the subsequent competition for consciousness and is included in the global broadcast. This coalition has entered consciousness. There is some evidence to indicate that this takes approximately 200-280ms from the beginning of a cycle for simple processing tasks, under the assumption that conscious perception involves synchronous oscillatory activity in brains [191].

Finally, an appropriate action is selected based on the contents of the conscious broadcast in the *action selection phase*. The schemes in Procedural Memory, in this case the two schemes representing the action to press the button and to release it, obtain activation based on how well the conscious contents match their context. A single action will then be selected in the Action Selection module. The chosen action is then passed to Sensory-Motor Memory where it is executed. When this occurs, the state of the button in the environment is set to the appropriate value by a snippet of code that could be called the LRT agent's "actuator."

The cognitive cycle durations of the LRT agent (283 ms averaged over 30 runs, see [69]) are comparable, but larger than the cycle durations inferred from the reaction times of adult humans (200ms according to [192]). This is consistent with recent neuroscientific evidence (e.g. [154, 193, 194], see also [69]) supporting the idea that single perception-action cycles may take longer than simple reaction time tasks under normal circumstances (e.g. more complex stimuli). We hypothesize that the main reason for humans being faster at such experiments is the effect of temporal expectation, which reduces reaction time (and has not yet been implemented in $LIDA_I$). A behavioral consequence of this is that reaction times to predictable stimuli are significantly lower than

[15] Embedding these processes into a cognitive cycle – which increases the duration of a single cycle – is one of the differences between LIDA and other cognitive architectures (see Section XI). Due to the early stage of the implemented $LIDA_I$ framework, this difference has not been empirically evaluated yet.

reactions to uncertain and temporally highly variable stimuli (see [195] for a review). From a neurophysiological point of view, increased activation levels can be observed in subcortical (the basal ganglia) and cortical action circuits (inferior parietal and premotor areas) prior to perceiving the stimulus. This increased activity is presumed to be capable of reducing the time required for action selection for predictable stimuli [196, 197], an effect not accounted for in this simulation.

**The LIDA$_I$ Allport agent**. This agent replicates a psychological experiment proposed by Allport [68], with the intention of comparing the *Discrete Moment Hypothesis* [198] with the *Continuous (Traveling) Moment Hypothesis*. The Discrete Moment Hypothesis posits that consciousness is comprised of distinct and non-overlapping conscious 'moments,' within which all time-order information is lost. In contrast, the *Continuous (Traveling) Moment Hypothesis* considers conscious 'moments' to correspond to continuously moving segments of the incoming sensory information. We used this experimental paradigm to show that LIDA's discrete consciousness position is an adequate model of human functional consciousness, despite Allport's conclusion that the *Discrete Moment Hypothesis* contradicts experimental evidence. Another goal of this simulation was to verify the timing parameters in a more complex setting [69].

In Allport's experiment, participants were seated in front of a screen, which displayed a single horizontal line, appearing in one of 12 positions on the screen (see Figure 4). This line rapidly changed position, moving upward. Upon reaching the topmost position, the screen was left blank for the same duration as the line took while traversing each of the 12 positions, and then the line appeared again on the bottom position. The cycle time ($\tau$) was controlled by the participant. At very large cycle times, participants could see the single line jumping from position to position. Upon decreasing $\tau$, they reported seeing multiple lines, moving together. At a specific cycle time S and below, participants reported seeing a stationary array of 12 lines flickering in synchrony. The participants had to arrive at the cycle time S, where they did not perceive any movement on the screen.

In separate trials participants first decreased the cycle time from a very high value (slow to fast), and then increased it from a very low value (fast to slow), at which all lines were seen simultaneously. Both times were recorded for each participant. These times were then compared to the predictions of the two hypotheses about consciousness. According to the *Discrete Moment Hypothesis*, there are two different cycle times $\tau$ at which all 12 lines are seen simultaneously on the screen and are perceived not to move. At $\tau_1$ = S, displaying all lines as well as the blank screen (left blank for S/2, the same time as the lines took to display) falls within one conscious 'moment'; thus subjects should not perceive any movement, since there will be no change between this conscious 'moment' (containing 12 lines and a blank screen) and the next one. At $\tau_2$ = S / 2, if the hypothesis of discrete conscious moments is accepted, no movement should be perceived either, since in this case conscious 'moments' containing all lines and the blank screen would alternate (in the first S/2 ms the 12 lines would be displayed, perceived simultaneously since they fall into one conscious 'moment'; and in the second S/2 ms there would be a blank screen – thus no moving lines could be perceived on the screen, just flickering). The cycle time at which subjects will perceive no movement will thus be S when decreasing $\tau$, and S/2 when increasing $\tau$. A significant difference between these two conditions is predicted.

In contrast, the *Continuous Moment Hypothesis* predicts that successive events are perceived to be simultaneous whenever, and as long as, they fall within the temporal constraints of the conscious 'moment.' Thus, since the criterion for determining S was not only momentary simultaneity but perpetual absence of perceived movement, there can be only one cycle time $\tau_1$=S at which this criterion is met (at $\tau_2$ = S / 2, the contents of a conscious 'moment' would change gradually from containing 12 lines to containing just the blank screen – thus there would be movement -, instead of just alternating between the two cases, as in the discrete case described above). There should be no difference between trials decreasing or increasing $\tau$. Allport [68] did not find significant differences between these two conditions, and thus argued against the *Discrete Moment Hypothesis*. However, despite LIDA$_I$'s consciousness mechanism being fundamentally discrete, we could successfully reproduce Allport's results with a LIDA$_I$-based agent.

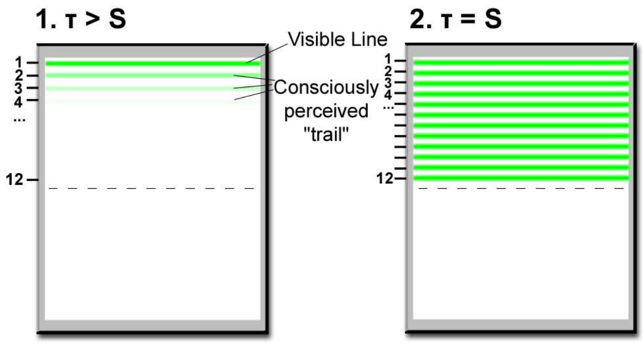


**Figure 4** The display and conscious percept in Allport's experiment. Lines were displayed in one of 12 positions, appearing to move upwards. Upon reaching the top, the screen was left blank for the same period as the lines required to traverse all 12 positions. $\tau$ denotes the total cycle time. At cycle times $\tau$ >S, subjects could see multiple lines moving together (left panel). At $\tau$=S, subjects saw all lines simultaneously and perceived no movement (right panel). (From [69] with permission)

The LIDA Allport agent was implemented similarly to the LRT agent. The major differences were the following: The Allport agent had a PAM consisting of twelve nodes, one for each line on the screen[16]. Feature detectors passed activation to these nodes depending on the line position. It also had two different schemes in the Procedural Memory. The first scheme became active when no movement was perceived on the screen, i.e. when the contents of multiple conscious broadcasts

[16] The twelve nodes, and the two schemes were hard-coded into this agent, since the implementations of perceptual and procedural learning in LIDA$_I$ are not yet finished.

contained all 12 lines; this scheme caused the agent to press the “no movement perceived” button. The second scheme was selected when a single line, or multiple lines, were perceived as moving by the agent, and resulted in the agent pressing the “movement perceived” button. For easier implementation, the agent did not change the cycle time (τ) itself – it only reacted to whether or not it perceived movement on the screen (the cycle time was changed gradually by the environment controller). There was only one cycle time at which the agent did not perceive movement, 96ms [69], which is consistent with Allport’s (1968) results (unlike the predictions of the *Discrete Moment Hypothesis*) and provides support for the claim that the temporal extent of a conscious “moment” of a LIDA agent is similar to that of a human. The main reason for this is that in the LIDA model, single conscious episodes are discrete but, contrary to Stroud’s [198] view, not necessarily distinct – a current conscious ‘moment’ can contain percepts from a previous moment. Whether or not an older percept remains conscious depends on how long in the past it has been perceived, and on attentional modulation, where percepts that are subjectively important and attended to can persist longer in consciousness.

**The LIDA Attentional agents**. We have developed two agents reproducing attention experiments to substantiate LIDA’s GWT-based attention and consciousness mechanism, the $\text{LIDA}_\text{I}$ Attention [23] and Attentional Blink [24] agents.

The first agent used an adapted version of Van Bockstaele’s experiment [199]. The environment consisted of a black screen with two white squares on the left and the right side of a central fixation cross, in which cues and the targets could appear, and which the agent had to respond to. After a fixation period, one of the white rectangles was randomly replaced by the cue (a colored rectangle) for 200ms, followed by the two white rectangles again for 20ms. Subsequently, the target (a small black rectangle) was randomly presented in one of the white rectangles until the agent responded, and the response time was measured. Humans [199] as well as the LIDA Attention agent [23] were faster by 20ms in congruent trials – in which the target appeared on the same side as the cue – than in incongruent trials (average response times were 360ms and 380ms). We hypothesize that the reason for the time difference is that in congruent trials, the procedural scheme responsible for the correct behavior has already been instantiated by the cue by the time the target arrives, and merely has to be selected and executed. In contrast, in incongruent trials the procedural scheme has to be instantiated as well as selected and executed; and this scheme instantiation takes an additional 20ms compared to the congruent case [23].

The second attentional agent reproduced the attentional blink (AB). The AB refers to the phenomenon of individuals often being unable to consciously perceive the second of two presented targets in a stream of target and distractor stimuli, if the second target T2 is presented within 200-500ms after the first target T1. A considerable number of effects have been discovered in AB experiments. It has been documented that the second target can be reported if presented immediately after T1 without a distractor in between (lag-1 sparing). Increasing T2 salience [200] or emotional arousal [201] also attenuates the AB effect. Although a large number of conceptual and computational models have been proposed, and the basic AB phenomenon is well understood, most models are unable to integrate and account for all phenomena and findings associated with the AB (see [202] or [200] for recent reviews of these models).

We have developed a LIDA-based attentional blink model [24] to computationally model visual AB experiment [203] and conceptually explain a wide range of phenomena. As can be seen from Figure 5, the $\text{LIDA}_\text{I}$ attentional blink agent was successful in reproducing the AB effect (using the data from [203]).

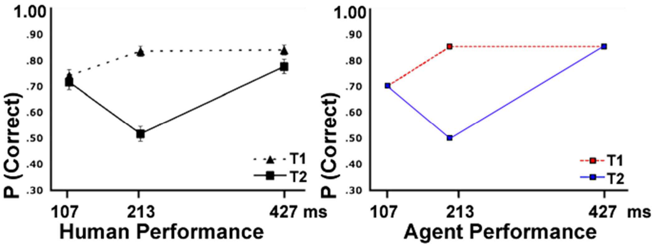


**Figure 5** The results for human subjects (panel B left, from[203]) and the LIDA AB agent (panel B right) (based on [24])

A LIDA-based AB model could provide a novel approach to understanding the AB for two reasons. First, LIDA is a general cognitive architecture, as opposed to specialized AB models, and integrates other cognitive phenomena as well. Thus, a larger range of phenomena can be accounted for (e.g. the attenuation by emotional arousal, or the AB attenuation in whole report experiments [204]). Although there is another cognitive architecture based AB model [205] based on ACT-R, this model is unable to conceptually explain the effects of emotional arousal, or the phenomenon of target confusion, because standard ACT-R does not include emotional processing or high-level vision [24]. $\text{LIDA}_\text{C}$ can account for both, although the former only on a conceptual level since emotions have not been implemented yet. Second, LIDA is also based on the GWT and thus provides a plausible account of attention and access consciousness, the most important mechanisms underlying the attentional blink.

## XI. Contrasting LIDA with Major Cognitive Architectures

A full description of other cognitive architectures would exceed the scope of this paper (see [4, 123, 206] for recent reviews), as would a comparison of LIDA’s features with the large number of other architectures (such comparisons can be found in [123, 207] or [208]). Instead, we will focus on a few significant differences between LIDA and existing major cognitive architectures, thereby highlighting how the LIDA model can complement research on cognitive architectures.

It is important to point out that the conceptual $\text{LIDA}_\text{C}$ model has only been partially computational implemented, and that reproduction of human data has only recently begun. Thus, it would be infeasible to compare $\text{LIDA}_\text{I}$’s simulation data to the wealth of data obtained from approximately 30 years of simulations of other architectures such as ACT-R. This is also the main reason why the present focus is both on computational as well as conceptual descriptions, both of which have proven highly useful in cognitive science by providing explanations and hypotheses guiding ongoing research [12, 16].

**Computational differences.** LIDA differs from many other cognitive architectures such as ACT-R [209], SOAR [10], CLARION [210, 211], EPIC [212], and Icarus [213] in a number of ways. The most significant of the implemented differences include:

*1. Adherence to grounded theories of cognition and lack of amodal symbols* (see Section II).

*2.Explicit and neuroscientifically plausible functional consciousness mechanism* (based on GWT) [22, 39, 40, 112] – the only other systems-level cognitive architecture explicitly addressing consciousness is CLARION, proposing a two-level representation which distinguishes between conscious and unconscious knowledge and processes [214]. The simulation of the Allport experiment is an example for the importance of functional consciousness for modeling human behavior.

*3. Specific explanation and subdivision of memory systems*. $LIDA_I$'s memory systems include sensory memory, working memory, perceptual associative memory, transient episodic memory, declarative memory, and procedural memory. $LIDA_C$ further includes sensory-motor memory, attentional memory [23] and spatial memory [215] which are currently under development. Transient episodic and declarative memories are modeled using a sparse-coded, noise-robust memory mechanism able to account for phenomena such as the tip-of-the-tongue effect or the remember-know distinction [216]. Although episodic memory is arguably an important part of human cognition, few cognitive models account for it [4]. SOAR has recently added an episodic memory module [217], in essence storing snapshots of the entire workspace content (short term declarative knowledge). In contrast, LIDA only stores events that appear in the conscious broadcast in its Transient Episodic Memory, and employs a modified instance of an SDM [92], resulting in a content-addressable, associative episodic memory with a decay rate measured in hours (consistent with Conway's [80] and Baddeley's [89] ideas).

*4. Sophisticated procedural memory and action selection*, facilitating high-level decision making and non-routine problem solving, which has been implemented [218], as well as procedural learning, which is a part of developmental learning that has been implemented [76] (but not yet integrated with the LIDA computational framework). The conceptual $LIDA_C$ model also includes other forms of developmental learning (see below). While the mentioned cognitive architectures address action selection and problem solving to different degrees [4], developmental learning is usually restricted to procedural learning. It should be mentioned that a number of architectures supporting developmental learning have been proposed in robotics (e.g. [219], see also [123]); however, these systems are usually not concerned with cognitive modeling.

*5. Complex, but detailed and effectual preconscious working memory* that enables binding and understanding (Section III and [79]). LIDA's workspace is fully grounded and modal [42], and consistent with Baddeley's theory of working memory [220].

**Conceptual differences**. There are also differences in the parts of the conceptual $LIDA_C$ model that have not been implemented computationally as of yet.

*1. The use of feelings and emotions as flexible motivators*. Emotions are not accounted for by the mentioned architectures, with the exception of SOAR, an architecture that has an implemented emotion model based on appraisal theory [95] and has emotionally motivated actions as well. The most important difference between SOAR's and $LIDA_C$'s emotion models is the derivation of affect intensity. SOAR employs an affect intensity model based on expected utility, whereas $LIDA_C$'s affect intensities are influenced by activations of nodes in the entire node structure (perceptual representation) representing an event.

*2. Several modes of human-like learning* [32, 221], including perceptual, spatial, procedural, attentional and episodic learning. None of the mentioned architectures account for all modes of learning mentioned in Section VI. Although conceptually developed, these ideas have not all been implemented computationally as of yet, except for episodic learning and procedural learning. Episodic learning [92] has been implemented in the $LIDA_I$ computational framework (Section IX), while procedural learning was implemented prior to the development of the framework [76] and is not yet part of it. Neither of these two have been validated by replicating human data as of yet. A LIDA-based agent implementing perceptual learning is now in progress.

As described above, we feel it is both useful and important to have fine-grained models of memory systems and learning mechanisms. The conceptual $LIDA_C$ model contains approaches and ideas of these processes, but few of these have been implemented in $LIDA_I$ to date. In contrast, many cognitive architectures have well-developed learning algorithms in specific domains. For example, reinforcement learning - a form of procedural learning - is implemented in ACT-R, SOAR and CLARION. The latter two are also able to form new semantic representations. SOAR also supports episodic learning (also implemented in LIDA). Despite their superiority in terms of implemented learning models, however, we believe that LIDA can still make a contribution here, since none of these architectures account for the entire variety of learning and memory suggested by the $LIDA_C$ model.

**Weaknesses of the LIDA model.** The major shortcomings of LIDA compared to other cognitive architectures include:

*1. No implementation of multicyclic cognition* (e.g. deliberation, volition, reasoning, …). This is the major strength of production-based systems (e.g, SOAR, ACT-R, …).

*2. No model of language*, as opposed to e.g. ACT-R [222] or SOAR [223, 224]. Current work simulating the learning of vervet monkey alarm calls (in preparation) is thought to be a precursor of adding language capability to LIDA.

*3. Very little work to date on causality and probability*[17], both of which would be important for real-world applications such as robotics. Example cognitive architectures soundly based on probability theory include Sigma [225] and Icarus [213].

*4. Lacking implementation of metacognition*, an important part of higher-level cognitive processing. In contrast, CLARION [226] includes implementations of meta-cognitive processes. (Metacognition was implemented in LIDA's predecessor IDA [227], but by means of adding a radically different architecture. Current plans are for an integrated inclusion of metacognition in $LIDA_c$.)

*5. Early stage of $LIDA_I$.* Although LIDA aims to be a general cognitive architecture with empirically verifiable predictions, only a small subset of $LIDA_C$ has been actually implemented and verified so far. LIDA is much more recent than many cognitive architectures – the first version of the computational framework was released less than two years ago [175].

**Comparison with similar systems.** The LIDA model is primarily an implementation and fleshing out of the Global Workspace Theory (GWT) of consciousness in cognition. GWT was inspired by the blackboard systems of Hayes-Roth and colleagues in AI [228], of McClelland in neural networks [229], and of Van Der Velde in cognitive sciences [230]. (see [231] for an overview.) Recall that a blackboard system, often dedicated to solving a complex ill-defined problem, consists of a family of knowledge sources (experts/agents), a shared blackboard data structure on which they write their suggested partial solutions, and a scheduling/control structure that regulates the writing to the blackboard. The LIDA model can be viewed as a blackboard system with its Workspace serving as the blackboard, each of the other memories and their processes acting as knowledge sources, and the attention/consciousness system constituting the scheduling/control structure. Note that unlike a typical blackboard system, LIDA models cognition, rather than being devoted to the solution of one problem.

There have also been other implementations of portions of GWT such as the Global Neuronal Workspace system of Dehaene and colleagues [232], Wallace's mathematical treatment of GWT [233], and Shanahan and colleagues' robotic GWT architecture [234]. None of these three constitutes a full, systems-level cognitive architecture.

## XII. How LIDA Addresses some of the Open Issues in Cognitive Architectures

In their recent review of research on cognitive architectures [4], Langley, Laird and Rogers list and discuss nine separate "open issues" which they suggest should drive current research on cognitive architectures. Here, we suggest that our LIDA architecture makes some contribution on six of those open issues. We will briefly describe each of those contributions in turn. Quotes in italics are from Langley et al, and serve to denote an open issue.

"*categorization and understanding*": The internal structure of LIDA's Workspace, including the CSM [79] (with its own internal structure) and the Conscious Contents Queue [235] are devoted precisely to the issue of low-level understanding of both external sensory input and internal state, and of the relation between them. The agents effecting this understanding are LIDA's structure-building codelets.

"*…architectures that directly support both episodic memory and reflective processes that operate on the structures it contains.*" The LIDA model includes both Transient Episodic Memory and Declarative Memory. Local associations from these memory systems form part of the content used by structure-building codelets to build new structures in the CSM in LIDA's Workspace.

"*… encode knowledge in a variety of formalisms, relate them to each other, and use them to support intelligent behavior more flexibly and effectively.*" LIDA employs distinct data structures, and distinct processes that operate on them for PAM, the episodic memories, and Procedural Memory. Their roles in the cognitive cycle relate them to one another, and allow them to support action selection.

"*…manage an agent's resources to selectively focus its perceptual attention, its effectors, and the tasks it pursues.*" LIDA's attentional mechanism (functional consciousness) performs just these functions.

"*…origin of agents' primary goals in terms of internal drives.*" Such internal drives are implemented explicated in $LIDA_C$ via feelings and emotions, providing flexibility in decision making.

"*…exhibit emotion in ways that link directly to other cognitive processes and that modulate intelligent behavior.*" Feelings and emotions play significant roles in each of $LIDA_C$'s major modules and processes, modulating action selection and learning as well.

## XIII. Concluding Remarks

We have provided a summary overview of the LIDA model, a systems-level conceptual and computational model of human cognition that is grounded in cognitive psychology and cognitive neuroscience, and whose architecture conceptually affords grounded cognition, attention, emotion, action selection, human-like learning, and other higher-level processes. We have also briefly described the LIDA computational Framework and have described simulations involving four LIDA software agents replicating psychological experiments and providing evidence that LIDA's cognitive cycle timing and LIDA's attention and consciousness mechanisms are comparable to human subjects. This is an important first step towards increasing LIDA's plausibility as a model of human cognition. More such replications are in progress. Furthermore, the LIDA cognitive architecture is suited not only for simulated experiments, but also for real-world applications. Its predecessor IDA was developed as a distribution agent for the Navy, communicating with sailors via email in natural language [66]. The abilities

[17] A paper on causality in LIDA is currently in preparation. Discussions on how best to incorporate probability in LIDA are ongoing.

and mechanisms required for this job can be used in a number of additional fields of application including artificial clerks, customer service agents, travel agents, loan officers in a bank, and many others [236]. Apart from human information agents, LIDA could also function on a physical robot (its reactivity facilitated by asynchronous operation and one-shot learning in the SDM). Work is underway to combine LIDA with the CRAM control system [237] and to embody it on a PR-2[18] humanoid robot.

Finally, we emphasize the importance of cognitive models such as LIDA. These models play a major role in cognitive science due to their usefulness in providing detailed and verifiable explanations for cognitive processes, and in providing hypotheses that can guide ongoing research [3, 4].

## Acknowledgment

The authors are profoundly indebted for ideas in this paper to too many past and present members of the Cognitive Computing Research Group at the University of Memphis and elsewhere to mention by name. Most recently they owe particular thanks to Bernard J. Baars and Ryan McCall.

## References

1. K. Lewin, *Field theory in sociall science: selected theoretical papers*, Harper & Row, 1951.
2. A. Newell, “You can’t play 20 questions with nature and win: Projective comments on the papers of this symposium,” *Visual information processing*, W. G. Chase, ed., Academic Press, 1973.
3. D.L. Hintzman, “Research Strategy in the Study of Memory: Fads, Fallacies, and the Search for the 'Coordinates of Truth',” *Perspectives on Psychological Science*, vol. 6, no. 3, 2011, pp. 253-271; DOI 10.1177/1745691611406924.
4. P. Langley, et al., “Cognitive Architectures: Research Issues and Challenges,” *Cogn Syst Res*, vol. 10, no. 2, 2009, pp. 141-160; DOI doi: 10.1016/j.cogsys.2006.07.004.
5. J.R. Anderson and C. Lebiere, “The Newell test for a theory of cognition,” *Behavioral and Brain Sciences*, vol. 26, no. 5, 2003, pp. 587-601.
6. S. Franklin and A.C. Graesser, “Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents,” *Intelligent Agents III*, Springer Verlag, 1997, pp. 21–35.
7. J. Fuster, “Upper processing stages of the perception-action cycle,” *Trends in Cognitive Science*, vol. 8, 2004, pp. 143-145.
8. W.J. Freeman, “The limbic action-perception cycle controlling goal-directed animal behavior,” *Neural Networks*, vol. 3, 2002, pp. 2249-2254.
9. S. Franklin, *Artificial Minds*, MIT Press, 1995.
10.John E. Laird, et al., “SOAR: An Architecture for General Intelligence,” *Artificial Intelligence*, vol. 33, 1987, pp. 1–64.
11.J.R. Anderson, *The Adaptive Character of Thought*, Erlbaum, 1990.
12.R. Sun, “The importance of cognitive architectures: An analysis based on CLARION,” *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 19, no. 2, 2007, pp. 159-193.
13.C.H. Anderson, et al., “Directed visual attention and the dynamic control of information flow,” *Neurobiology of Attention*, L. Itti, et al., eds., Academic Press/Elsevier, 2005, pp. 11–17.
14.A. Baddeley and L. Weiskrantz, *Attention: Selection, Awareness, & Control*, Oxford University Press, 1993.
15.J. Tsotsos and N. Bruce, “Computational foundations for attentive processes,” *Scholarpedia*, vol. 3, no. 12, 2008, pp. 6545-6545.
16.J.R. Anderson, et al., “ACT-R: A theory of higher level cognition and its relation to visual attention,” *Human Computer Interaction*, vol. 12, no. 4, 1997, pp. 439-462.
17.M.W. Spratling and M.H. Johnson, “A feedback model of visual attention,” *J Cogn Neurosci*, vol. 16, no. 2, 2004, pp. 219-237; DOI 10.1162/089892904322984526.
18.Dario D. Salvucci, “An integrated model of eye movements and visual encoding,” *Cogn Syst Res*, vol. 1, 2001, pp. 201–220.
19.S.-n. Yang, “An oculomotor-based model of eye movements in reading: The competition/interaction model,” *Cogn Syst Res*, vol. 7, no. 1, 2006, pp. 56-69; DOI 10.1016/j.cogsys.2005.07.005

20.E.M. Altmann and W.D. Gray, “An integrated model of cognitive control in task switching,” *Psychological Review*, vol. 115, 2008, pp. 602-639.
21.T. Zhang, et al., “Characterisation of mental models in a virtual reality-based multitasking scenario using measures of situation awareness ” *Theoretical Issues in Ergonomics Science*, vol. 11, no. 1 & 2, 2010, pp. 99 - 118.
22.Bernard J. Baars, *A Cognitive Theory of Consciousness*, Cambridge University Press, 1988.
23.U. Faghihi, et al., “A computational model of attentional learning in a cognitive agent,” *Biologically Inspired Cognitive Architectures*, 2012; DOI 10.1016/j.bica.2012.07.003.
24.T. Madl and S. Franklin, “A LIDA-based Model of the Attentional Blink,” *Proc. International Conference on Cognitive Modelling*, 2012.
25.Gerald M. Edelman, *Neural Darwinism*, Basic Books, 1987.
26.J.-C.D. Sandoz, Jean-Marc
Giurfa, Martin, “Invertebrate Learning and Memory,” *Book Invertebrate Learning and Memory*, Series Invertebrate Learning and Memory, ed., Frontiers E-books, 2011, pp.
27.R.C. Atkinson and R.M. Shiffrin, “Human memory: A proposed system and its control processes,” *The psychology of learning and motivation*, vol. 2, 1968, pp. 89-195.
28.D.F. Sherry and D.L. Schacter, “The evolution of multiple memory systems,” *Psychological review*, vol. 94, no. 4, 1987, pp. 439.
29.L.R. Squire, “Memory systems of the brain: a brief history and current perspective,” *Neurobiology of learning and memory*, vol. 82, no. 3, 2004, pp. 171-177.
30.S. Franklin, “Deliberation and Voluntary Action in ‘Conscious’ Software Agents,” *Neural Network World*, vol. 10, 2000, pp. 505–521
31.S. Franklin, et al., “The Role of Consciousness in Memory,” *Brains, Minds and Media*, vol. 1, 2005, pp. 1–38, pdf.
32.S. Franklin and F.G.J. Patterson, “The LIDA Architecture: Adding New Modes of Learning to an Intelligent, Autonomous, Software Agent,” *IDPT-2006 Proceedings (Integrated Design and Process Technology)*, Society for Design and Process Science, 2006.
33.S. Franklin and U. Ramamurthy, “Motivations, Values and Emotions: Three sides of the same coin,” *Proceedings of the Sixth International Workshop on Epigenetic Robotics*128, Lund University Cognitive Studies, 2006, pp. 41–48.
34.A. Negatu, “Cognitively Inspired Decision Making for Software Agents: Integrated Mechanisms for Action Selection, Expectation, Automatization and Non-Routine Problem Solving: Ph.D,” Dissertation, The University of Memphis, Memphis TN USA, 2006.
35.Sidney K. D'Mello, et al., “A Procedural Learning Mechanism for Novel Skill Acquisition,” *Proceeding of Adaptation in Artificial and Biological Systems, AISB'06*1, T. Kovacs and James A. R. Marshall, eds., Society for the Study of Artificial Intelligence and the Simulation of Behaviour, 2006, pp. 184–185.
36.A. Negatu and S. Franklin, “An action selection mechanism for 'conscious' software agents,” *Cognitive Science Quarterly*, vol. 2, no. special issue on "Desires, goals, intentions, and values: Computational architectures." Guest editors Maria Miceli and Cristiano Castelfranchi., 2002, pp. 363–386.
37.U. Ramamurthy, et al., “LIDA: A Working Model of Cognition,” *Proceedings of the 7th International Conference on Cognitive Modeling*, D. Fum, et al., eds., Edizioni Goliardiche, 2006, pp. 244–249.
38.U. Ramamurthy, et al., “LIDA: A Computational Model of Global Workspace Theory and Developmental Learning,” *Proc. BICS 2006: Brain Inspired Cognitive Systems*, 2006.
39.B. Baars, “In the theatre of consciousness. Global Workspace Theory, a rigorous scientific theory of consciousness,” *Journal of Consciousness Studies*, vol. 4, 1997, pp. 292–309.
40.Bernard J. Baars, “The conscious access hypothesis: origins and recent evidence,” *Trends in Cognitive Science*, vol. 6, 2002, pp. 47–52.
41.B.J. Baars, “Global workspace theory of consciousness: toward a cognitive neuroscience of human experience,” *Prog Brain Res*, vol. 150, 2005, pp. 45-53; DOI 10.1016/S0079-6123(05)50004-9.
42.L.W. Barsalou, “Grounded Cognition,” *Annual Review of Psychology*, vol. 59, 2008, pp. 617–645.
43.F.J. Varela, et al., *The Embodied Mind*, MIT Press., 1991.

[18] http://www.willowgarage.com/pages/pr2

44.L.W. Barsalou, "Perceptual symbol systems," *Behavioral and Brain Sciences*, vol. 22, 1999, pp. 577–609.
45.A.D. Baddeley and G.J. Hitch, "Working memory," *The Psychology of Learning and Motivation*, G. A. Bower, ed., Academic Press, 1974, pp. 47–89.
46.A.D. Baddeley, "The episodic buffer: a new component of working memory?," *Trends in Cognitive Science*, vol. 4, 2000, pp. 417–423.
47.A.M. Glenberg, "What memory is for," *Behavioral and Brain Sciences*, vol. 20, 1997, pp. 1–19.
48.K.A. Ericsson and W. Kintsch, "Long-term working memory," *Psychological Review*, vol. 102, 1995, pp. 211–245.
49.A. Sloman, "What Sort of Architecture is Required for a Human-like Agent?," *Foundations of Rational Agency*, M. Wooldridge and A. S. Rao, eds., Kluwer Academic Publishers, 1999, pp. 35–52.
50.D.R. Hofstadter and M. Mitchell, "The Copycat Project: A model of mental fluidity and analogy-making," *Advances in connectionist and neural computation theory, Vol. 2: logical connections*, K. J. Holyoak and J. Barnden, eds., Ablex, 1995, pp. 205–267.
51.J. Marshall, "Metacat: a self-watching cognitive architecture for analogy-making," *Proc. 24th Annual Conference of the Cognitive Science Society*, 2002, pp. 631-636.
52.P. Kanerva, *Sparse Distributed Memory*, The MIT Press, 1988.
53.R.P.N. Rao and O. Fuentes, "Hierarchical Learning of Navigational Behaviors in an Autonomous Robot using a Predictive Sparse Distributed Memory," *Machine Learning*, vol. 31, 1998, pp. 87-113.
54.Gary L. Drescher, *Made-Up Minds: A Constructivist Approach to Artificial Intelligence*, MIT Press, 1991, p. 236.
55.H.H. Chaput, et al., "Constructivist Learning: A Neural Implementation of the Schema Mechanism," *Proc. Proceedings of WSOM '03: Workshop for Self-Organizing Maps*, 2003.
56.P. Maes, "How to do the right thing," *Connection Science*, vol. 1, 1989, pp. 291–323.
57.T. Tyrrell, "An Evaluation of Maes's Bottom-Up Mechanism for Behavior Selection," *Adaptive Behavior*, vol. 2, 1994, pp. 307-348.
58.Rodney A. Brooks, "How to build complete creatures rather than isolated cognitive simulators," *Architectures for Intelligence*, K. VanLehn, ed., Lawrence Erlbaum Associates, 1991, pp. 225–239.
59.T. Ziemke, "Embodied AI as science: Models of embodied cognition, embodied models of cognition, or both?," *Lect Notes Artif Int*, vol. 3139, 2004, pp. 27-36.
60.S. Franklin, "Autonomous Agents as Embodied AI," *Cybernetics and Systems*, vol. 28, 1997, pp. 499–520.
61.M. Wilson, "Six views of embodied cognition," *Psychonomic Bulletin & Review*, vol. 9, no. 4, 2002, pp. 625 - 636.
62.W. Schwarz and I.M. Keus, "Moving the eyes along the mental number line: Comparing SNARC effects with saccadic and manual responses," *Attention, Perception, & Psychophysics*, vol. 66, no. 4, 2004, pp. 651-664.
63.T. Hansen, et al., "Memory modulates color appearance," *Nat Neurosci*, vol. 9, 2006, pp. 1367–1368.
64.D. Vernon, et al., *A roadmap for cognitive development in humanoid robots*, Springer Verlag, 2010.
65.D.P. Benjamin, et al., "Embodying a cognitive model in a mobile robot," 2006, pp. 638407-638407.
66.S. Franklin, "IDA: A Conscious Artifact?," *Journal of Consciousness Studies*, vol. 10, 2003, pp. 47–66.
67.B.J. Baars and S. Franklin, "An architectural model of conscious and unconscious brain functions: Global Workspace Theory and IDA," *Neural Networks*, vol. 20, 2007, pp. 955-961.
68.D.A. Allport, "Phenomenal simultaneity and the perceptual moment hypothesis," *British Journal of Psychology*, vol. 59, 1968, pp. 395-406.
69.T. Madl, et al., "The Timing of the Cognitive Cycle," *PLoS ONE*, vol. 6, no. 4, 2011, pp. e14803.
70.A. Newell, *Unified Theories of Cognition*, Harvard University Press, 1990.
71.B. Merker, "The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution," *Consciousness and Cognition*, vol. 14, 2005, pp. 89–114.
72.S. Franklin, "Evolutionary Pressures and a Stable World for Animals and Robots: A Commentary on Merker," *Consciousness and Cognition*, vol. 14, 2005, pp. 115–118.
73.M. Sigman and S. Dehaene, "Dynamics of the Central Bottleneck: Dual-Task and Task Uncertainty," *PLoS Biol.*, vol. 4, 2006.
74.N. Uchida, et al., "Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making," *Nature Reviews Neuroscience*, vol. 7, 2006, pp. 485–491.
75.A. Newell and H.A. Simon, "Computer science as empirical inquiry: symbols and search," *Commun. ACM*, vol. 19, no. 3, 1976, pp. 113-126; DOI 10.1145/360018.360022.
76.G. Colman, "Procedural Memory Construction V2 (working document)," *Book Procedural Memory Construction V2 (working document)*, Series Procedural Memory Construction V2 (working document), ed., University of Memphis, 2007, pp.
77.J.Â.M. Fuster, "Physiology of executive functions: The perception-action cycle," 2002.
78.Bernard J. Baars and S. Franklin, "How conscious experience and working memory interact," *Trends in Cognitive Science*, vol. 7, 2003, pp. 166–172.
79.R. McCall, et al., "Grounded Event-Based and Modal Representations for Objects, Relations, Beliefs, Etc," *FLAIRS-23, Daytona Beach, FL*, 2010.
80.Martin A. Conway, "Sensory–perceptual episodic memory and its context: autobiographical memory," *Philos. Trans. R. Soc. Lond B.*, vol. 356, 2001, pp. 1375–1384.
81.J. Snaider, et al., "Time production and representation in a conceptual and computational cognitive model," *Cognitive Systems Research*, vol. In Press, Corrected Proof, 2011; DOI 10.1016/j.cogsys.2010.10.004.
82.A. Negatu, et al., "Automatization for Software Agents," in review.
83.A. Negatu, et al., "A non-routine problem solving mechanism for a general cognitive agent architecture," *Problem Solving: Techniques, Steps, and Processes*, Nova Science Publishers, 2006.
84.A.S. Negatu, "Cognitively Inspired Decision Making for Software Agents: Integrated Mechanisms for Action Selection, Expectation, Automatization and Non-Routine Problem Solving," *The Knowledge Engineering Review*, vol. 24, no. 04, 2009, pp. 410-410.
85.W. James, *The Principles of Psychology*, Harvard University Press, 1890.
86.Y.K. Shin, et al., "A review of contemporary ideomotor theory," *Psychological Bulletin*, vol. 136, no. 6, 2010, pp. 943-974; DOI 10.1037/a0020541.
87.E. Berne, *Games People Play – The Basic Hand Book of Transactional Analysis*, Ballantine Books, 1964.
88.S. Franklin, et al., "The Role of Consciousness in Memory," *Brains, Minds and Media*, vol. 1, no. bmm150 (urn:nbn:de:0009-3-1505), 2005.
89.A. Baddeley, et al., *Episodic Memory*, Oxford University Press, 2001, p. 294.
90.L. Standing, "Learning 10,000 pictures," *Quaterly Journal of Experimental Psychology*, vol. 25, 1973, pp. 207–222.
91.H.P. Bahrick, "Semantic memory content in permastore: fifty years of memory for Spanish learned in school," *J. Exp. Psychol. Gen*., vol. 113, 1984, pp. 1–29.
92.U. Ramamurthy, et al., "Modified Sparse Distributed Memory as Transient Episodic Memory for Cognitive Software Agents," *Proceedings of the International Conference on Sstems, Man and Cybernetics*, IEEE, 2004.
93.A. Morse, et al., "Towards an enactive cognitive architecture," 2008.
94.R.P. Marinier, et al., "A computational unification of cognitive behavior and emotion," *Cogn Syst Res*, vol. 10, no. 1, 2009, pp. 48-69; DOI DOI 10.1016/j.cogsys.2008.03.004.
95.R. Marinier and J. Laird, "Toward a comprehensive computational model of emotions and feelings," 2004.
96.E. Hudlicka, "Beyond cognition: Modeling emotion in cognitive architectures," *Proceedings of the Sixth International Conference on Cognitive Modeling*, 2004, pp. 118-123.
97.J. Gratch, et al., "Modeling the cognitive antecedents and consequences of emotion," *Cogn Syst Res*, vol. 10, no. 1, 2009, pp. 1-5; DOI DOI 10.1016/j.cogsys.2008.06.001.
98.S. Marsella, et al., *Computational Models of Emotion. Blueprint for Affective Computing. KR Scherer, T. Banziger and E. Roesch*, Oxford University Press, 2010.
99.Victor S. Johnston, *Why We Feel:The Science of Human Emotions*, Perseus Books, 1999, p. 210.
100. K.R. Scherer, "Appraisal considered as a process of multilevel sequential checking," *Appraisal processes in emotion: Theory, methods, research*, vol. 92, 2001, pp. 120-120.
101. J. Russell, "Core affect and the psychological construction of emotion," *Psychological Review*, vol. 110, 2003, pp. 145-172.
102. R. Lazarus, *Emotion and adaptation*, Oxford University Press, 1991.
103. A. Ortony, et al., *The cognitive structure of emotions*, Cambridge University Press., 1988.
104. K.R. Scherer, "The dynamic architecture of emotion: Evidence for the component process model," *Cognition & Emotion*, vol. 23, no. 7, 2009, pp. 1307-1351; DOI 10.1080/02699930902928969.
105. G. Bower, "Mood and memory," *American Psychologist*, vol. 36, 1981, pp. 129-148.

106. C.A. Smith and L.D. Kirby, “Toward delivering on the promise of appraisal theory,” *Appraisal processes in emotion: Theory, Methods, Research*, K. R. Scherer, et al., eds., Oxford University Press, 2001, pp. 121-138.
107. D. Friedlander and S. Franklin, “LIDA and a Theory of Mind,” *Artificial General Intelligence 2008*, P. Wang, et al., eds., IOS Press, 2008, pp. 137-148.
108. W. Wallach, et al., “Consciousness and Ethics: Artificially Conscious Moral Agents,” *International Journal of Machine Consciousness*, vol. 3, no. 1, 2011, pp. 177-192
109. R.J. Thompson, et al., “Concurrent and prospective relations between attention to emotion and affect intensity: An experience sampling study,” *Emotion*, vol. 11, no. 6, 2011, pp. 1489-1489.
110. P.A. Palmieri, et al., “Measuring clarity of and attention to emotions,” *Journal of personality assessment*, vol. 91, no. 6, 2009, pp. 560-567.
111. J. Gratch and S. Marsella, “Evaluating a computational model of emotion,” *Autonomous Agents and Multi-Agent Systems*, vol. 11, no. 1, 2005, pp. 23-43.
112. B.J. Baars and S. Franklin, “Consciousness is computational: The LIDA model of Global Workspace Theory.,” *International Journal of Machine Consciousness*, vol. 1, no. 1, 2009, pp. 23-32.
113. R.H. Luchsinger, et al., “Bacterial swimming strategies and turbulence,” *Biophysical journal*, vol. 77, no. 5, 1999, pp. 2377-2386.
114. T.R. Johnson, “Control in Act-R and Soar,” *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*, M. Shafto and P. Langley, eds., Lawrence Erlbaum Associates, 1997, pp. 343-348.
115. W. Wallach, et al., “A Conceptual and Computational Model of Moral Decision Making in Human and Artificial Agents,” *Topics in Cognitive Science, special issue on Cognitive Based Theories of Moral Decision Making*, W. Wallach and S. Franklin, eds., Cognitive Science Society, 2010, pp. 454-485.
116. R.M. Yerkes and J.D. Dodson, “The Relationship of Strength of Stimulus to Rapidity of Habit Formation,” *J Comp Neurol Psycho*, vol. 18, 1908, pp. 459-482.
117. J. Zhu and P. Thagard, “Emotion and action,” *Philosophical Psychology*, vol. 15, 2002, pp. 19–36.
118. C. Skarda and Walter J. Freeman, “How Brains Make Chaos in Order to Make Sense of the World,” *Behavioral and Brain Sciences*, vol. 10, 1987, pp. 161–195.
119. J.A.S. Kelso, *Dynamic Patterns: The Self Organization of Brain and Behavior*, MIT Press, 1995.
120. G. Buzsaki, *Rhythms of the Brain*, Oxford University Press, 2006.
121. J.P. Spencer and G. Schöner, “Bridging the representational gap in the dynamic systems approach to development,” *Developmental Science*, vol. 6, no. 4, 2003, pp. 392-412; DOI 10.1111/1467-7687.00295.
122. G. Schöner, “Dynamical systems approaches to cognition,” *Toward a Unified Theory of Development: Connectionism and Dynamic Systems Theory Re-Considered*, J. P. Spencer, et al., eds., Oxford University Press., 2008.
123. B. Goertzel, et al., “World survey of artificial brains, Part II: Biologically inspired cognitive architectures,” *Neurocomputing*, 2010.
124. W. Erlhagen, et al., “The distribution of neuronal population activation (DPA) as a tool to study interaction and integration in cortical representations,” *Journal of Neuroscience Methods*, vol. 94, no. 1, 1999, pp. 53-66; DOI 10.1016/s0165-0270(99)00125-9.
125. D. Jancke, et al., “Parametric Population Representation of Retinal Location: Neuronal Interaction Dynamics in Cat Primary Visual Cortex,” *The Journal of Neuroscience*, vol. 19, no. 20, 1999, pp. 9016-9028.
126. J.P. Spencer, et al., “Moving toward a grand theory while valuing the importance of the initial conditions,” *Toward a unified theory of development: Connectionism and dynamic systems theory re-considered*, J. P. Spencer, et al., eds., Oxford University Press, 2009, pp. 354-372.
127. R.H. Cuijpers and W. Erlhagen, “Implementing Bayes' rule with neural fields,” *Artificial Neural Networks - Icann 2008, Pt Ii*, vol. 5164, 2008, pp. 228-237.
128. S.I. Amari, “Dynamics of Pattern Formation in Lateral-Inhibition Type Neural Fields,” *Biol Cybern*, vol. 27, no. 2, 1977, pp. 77-87.
129. A. Bastian, et al., “Preshaping and continuous evolution of motor cortical representations during movement preparation,” *Eur J Neurosci*, vol. 18, no. 7, 2003, pp. 2047-2058; DOI 10.1046/j.1460-9568.2003.02906.x.
130. Y.Z. Chen, et al., “Optimal decoding of correlated neural population responses in the primate visual cortex,” *Nat Neurosci*, vol. 9, no. 11, 2006, pp. 1412-1420; DOI Doi 10.1038/Nn1792.
131. R. McCall, et al., “Sensory and Perceptual Scene Representation
”, In review.
132. J.S. Johnson, et al., “Moving to higher ground: The dynamic field theory and the dynamics of visual cognition,” *New Ideas in Psychology*, vol. 26, no. 2, 2008, pp. 227-251; DOI 10.1016/j.newideapsych.2007.07.007.
133. E. Bicho, et al., “A dynamic field approach to goal inference, error detection and anticipatory action selection in human-robot collaboration,” *New Frontiers in Human-Robot Interaction*, K. Dautenhah and J. Saunders, eds., John Benjamins Publishing Company, 2011, pp. 135-164.
134. W. Erlhagen and G. Schoner, “Dynamic field theory of movement preparation,” *Psychological Review*, vol. 109, no. 3, 2002, pp. 545-572; DOI Doi 10.1037//0033-295x.109.3.545.
135. C. Alain, et al., “A distributed cortical network for auditory sensory memory in humans,” *Brain Res.*, vol. 812, 1998, pp. 23–37.
136. S. Magnussen, “Low-level memory processes in vision,” *Trends in Neurosciences*, vol. 23, no. 6, 2000, pp. 247-251; DOI 10.1016/s0166-2236(00)01569-1.
137. J.M. Fuster, “The cognit: a network model of cortical representation,” *International Journal of Psychophysiology*, vol. 60, no. 2, 2006, pp. 125-132.
138. J.M. Fuster, “Cortex and memory: Emergence of a new paradigm,” *Journal of cognitive neuroscience*, vol. 21, no. 11, 2009, pp. 2047-2072.
139. L. Davachi, et al., “Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories,” *Proc. Natl. Acad. Sci. USA*, vol. 100, 2003, pp. 2157–2162.
140. B.D. Winters and T.J. Bussey, “Transient inactivation of perirhinal cortex disrupts encoding, retrieval, and consolidation of object recognition memory,” *J Neurosci*, vol. 25, no. 1, 2005, pp. 52-61; DOI 25/1/52 [pii] 10.1523/JNEUROSCI.3827-04.2005.
141. John G. Taylor and Nickolaos F. Fragopanagos, “The interaction of attention and emotion,” *Neural Networks*, vol. 18, 2005, pp. 353–369.
142. Alex C. Keene, et al., “Drosophila Dorsal Paired Medial Neurons Provide a General Mechanism for Memory Consolidation,” *Current Biology*, vol. 16, 2006, pp. 1524–1530.
143. H. Kim, et al., “Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain,” *PLoS Biol.*, vol. 4, 2006, pp. e233.
144. A.D. Baddeley, “The concept of episodic memory,” *Episodic Memory*, A. D. Baddeley, et al., eds., Oxford University Press, 2001, pp. 1–10.
145. H.J. Markowitsch, “Neuroanatomy of memory,” *The Oxford handbook of memory*, E. Tulving and F. I. M. Craik, eds., Oxford University Press, 2000, pp. 465–484.
146. L. Shastri, “Episodic memory and cortico-hippocampal interactions,” *Trends in Cognitive Sciences*, vol. 6, 2002, pp. 162–168.
147. J. Ferbinteanu and Matthew L. Shapiro, “Prospective and Retrospective Memory Coding in the Hippocampus,” *Neuron*, vol. 40, 2003, pp. 1227–1239.
148. Martin A. Conway, “Sensory-perceptual episodic memory and its context: autobiographical memory,” *Episodic Memory*, A. Baddeley, et al., eds., Oxford University Press, 2002, pp. 53–70.
149. C.M. Bird and N. Burgess, “The hippocampus and memory: insights from spatial processing,” *Nature Reviews Neuroscience*, vol. 9, no. 3, 2008, pp. 182-194; DOI 10.1038/nrn2335.
150. N. Burgess, “Spatial cognition and the brain,” *Annals of the New York Academy of Sciences*, vol. 1124, 2008, pp. 77-97; DOI 10.1196/annals.1440.002.
151. S. Dehaene and J.P. Changeaux, “Neural mechanisms for access to consciousness,” *The cognitive neurosciences*3, M. Gazzaniga, ed., MIT Press, 2004, pp. 1145-1157.
152. Gerald M. Edelman and G. Tononi, *A Universe of Consciousness*, Basic Books, 2000.
153. R. Gaillard, et al., “Converging intracranial markers of conscious access,” *PLoS Biology*, vol. 7, no. 3, 2009, pp. e1000061; DOI 10.1371/journal.pbio.1000061
154. S. Doesburg, et al., “Rhythms of Consciousness: Binocular Rivalry Reveals Large-Scale Oscillatory Network Dynamics
Mediating Visual Perception,” *PLoS ONE*, vol. 4, no. 7, 2009, pp. e6142; DOI 10.1371/journal.pone.0006142.
155. Steven W. Kennerley, et al., “Optimal decision making and the anterior cingulate cortex,” *Nat Neurosci*, vol. 9, 2006, pp. 940–947.
156. C. Pittenger, et al., “Impaired bidirectional synaptic plasticity and procedural memory formation in striatum-specific cAMP response element-binding protein-deficient mice,” *J. Neurosci.*, vol. 26, 2006, pp. 2808–2813.
157. B. Pasquereau, et al., “Shaping of motor responses by incentive values through the basal ganglia,” *Journal of Neuroscience*, vol. 27, 2007, pp. 1176-1183.
158. K. MacDonald, “Effortful Control, Explicit Processing and the Regulation of Human Evolved Predispositions,” *Psychological Review*, vol. 115, no. 4, 2008, pp. 012–1031.

159. J. Rowe, et al., “Rule-Selection and Action-Selection have a Shared Neuroanatomical Basis in the Human Prefrontal and Parietal Cortex,” *Cerebral Cortex*, vol. 18, 2008, pp. 2275--2285; DOI 10.1093/cercor/bhm249.
160. A. Del Cul, et al., “Causal role of prefrontal cortex in the threshold for access to consciousness,” *Brain*, vol. 132, no. Pt 9, 2009, pp. 2531-2540; DOI 10.1093/brain/awp111.
161. Hisham E. Atallah, et al., “Separate neural substrates for skill learning and performance in the ventral and dorsal striatum,” *Nat Neurosci*, vol. 10, 2006, pp. 126–131.
162. B. Horwitz and M.F. Glabus, “Neural modeling and functional brain imaging: the interplay between the data-fitting and simulation approaches,” *International review of neurobiology*, vol. 66, 2005, pp. 267-290.
163. A. Stocco and J.R. Anderson, “Endogenous control and task representation: an fMRI study in algebraic problem-solving,” *Journal of cognitive neuroscience*, vol. 20, no. 7, 2008, pp. 1300-1314.
164. A. Terao, et al., “An fMRI study of the Interplay of Symbolic and Visuo-spatial Systems inMathematical Reasoning,” *Human-Computer Interaction Institute*, 2004, pp. 2-2.
165. J.R. Anderson, “Using brain imaging to guide the development of a cognitive architecture,” *Integrated models of cognitive systems*, 2007, pp. 49-62.
166. J.R. Anderson, et al., “A central circuit of the mind,” *Trends in cognitive sciences*, vol. 12, no. 4, 2008, pp. 136-143.
167. D. Lloyd, “Virtual lesions and the not-so-modular brain,” *Journal of the International Neuropsychological Society*, vol. 6, no. 5, 2000, pp. 627-635.
168. G.C. Van Orden and K.R. Paap, “Functional neuroimages fail to discover pieces of mind in the parts of the brain,” *Philosophy of Science*, 1997, pp. 85-94.
169. J.W. Lewis and D.C. Van Essen, “Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey,” *The Journal of comparative neurology*, vol. 428, no. 1, 2000, pp. 112-137.
170. O. Sporns and R. Kotter, “Motifs in brain networks,” *PLoS Biol*, vol. 2, no. 11, 2004, pp. e369; DOI 10.1371/journal.pbio.0020369.
171. M. Shanahan, *Embodiment and the Inner Life*, Oxford University Press, 2010.
172. G. Buzsáki, *Rhythms of the Brain*, Oxford University Press, USA, 2006.
173. S.F. Strain, et al., “Brain rhythms, cognitive cycles and mental moments,” in preparation.
174. S. Franklin, et al., “Global Workspace Theory, its LIDA Model and the Underlying Neuroscience,” *Biologically Inspired Cognitive Architectures*, 2012.
175. J. Snaider, et al., “The LIDA Framework as a General Tool for AGI,” *Proc. The Fourth Conference on Artificial General Intelligence (Springer Lecture Notes in Artificial Intelligence)*, Springer, 2011.
176. E. Gamma, *Design patterns : elements of reusable object-oriented software*, Addison-Wesley, 1995, p. xv, 395 p.
177. J. Snaider, et al., “The LIDA Framework as a General Tool for AGI,” *Book The LIDA Framework as a General Tool for AGI*, Series The LIDA Framework as a General Tool for AGI, ed., 2011, pp.
178. P. Kanerva, “Hyperdimensional Computing: An Introduction to computing in distributed representation with high-dimensional random vectors,” *Cognitive Computation*, vol. 1, no. 2, 2009, pp. 139-159.
179. G.E. Hinton, “Mapping part-whole hierarchies into connectionist networks,” *Artificial Intelligence*, vol. , no. 46, 1990, pp. 47-75.
180. T.A. Plate, *Holographic Reduced Representation: distributed representation of cognitive structure*, CSLI, 2003.
181. J. Snaider, “Integer Sparse Distributed Memory and Modular Composite Representation,” University of Memphis, Memphis, 2012.
182. J. Snaider and S. Franklin, “Modular Composite Representation,” in review.
183. J. Snaider, et al., “Integer sparse distributed memory: Analysis and results,” *Neural Networks*, 2013; DOI http://dx.doi.org/10.1016/j.neunet.2013.05.005.
184. J. Snaider and S. Franklin, “Extended Sparse Distributed Memory and Sequence Storage,” *Cognitive Computation*, vol. 4, no. 2, 2012, pp. 172-180.
185. J. Snaider and S. Franklin, “Extended Sparse Distributed Memory,” *Proc. Biological Inspired Cognitive Architectures 2011*, 2011.
186. X. Huang and M.A. Paradiso, “V1 response timing and surface filling-in,” *Journal of Neurophysiology*, vol. 100, no. 1, 2008, pp. 539-547; DOI 10.1152/jn.00997.2007.
187. H. Kirchner, et al., “Ultra-rapid sensory responses in the human frontal eye field region,” *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 29, no. 23, 2009, pp. 7599-7606; DOI 10.1523/JNEUROSCI.1233-09.2009.
188. A. Raffone and N. Srinivasan, “An adaptive workspace hypothesis about the neural correlates of consciousness: insights from neuroscience and meditation studies,” *Progress in Brain Research*, vol. 176, 2009, pp. 161-180.
189. E. Thompson and F.J. Varela, “Radical embodiment: neural dynamics and consciousness,” *Trends in Cognitive Sciences*, vol. 5, no. 10, 2001, pp. 418-425.
190. H. Liu, et al., “Timing, Timing, Timing: Fast Decoding of Object Information from Intracranial Field Potentials in Human Visual Cortex,” *Neuron*, vol. 62, no. 2,, 2009, pp. 281-290; DOI 10.1016/j.neuron.2009.02.025.
191. S.M. Doesburg, et al., “Large-Scale Gamma-Band Phase Synchronization and Selective Attention,” *Cerebral Cortex*, vol. 18, no. 2, 2007, pp. 386-396; DOI 10.1093/cercor/bhm073.
192. A.T. Welford and J.M.T. Brebner, *Reaction times*, Academic Pr., 1980.
193. S.M. Doesburg, et al., “Large-scale gamma-band phase synchronization and selective attention,” *Cereb Cortex*, vol. 18, no. 2, 2008, pp. 386-396; DOI bhm073 [pii]
10.1093/cercor/bhm073.
194. T.W. Picton, “The P300 wave of the human event-related potential,” *J Clin Neurophysiol*, vol. 9, no. 4, 1992, pp. 456-479.
195. P. Niemi and R. Näätänen, “Foreperiod and simple reaction time,” *Psychological Bulletin*, vol. 89, no. 1, 1981, pp. 133-162.
196. J. Coull and A. Nobre, “Dissociating explicit timing from temporal expectation with fMRI,” *Current Opinion in Neurobiology*, vol. 18, no. 2, 2008, pp. 137-144; DOI 10.1016/j.conb.2008.07.011.
197. A. Nobre, et al., “The hazards of time,” *Current Opinion in Neurobiology*, vol. 17, no. 4, 2007, pp. 465-470; DOI 10.1016/j.conb.2007.07.006.
198. J.M. Stroud, “The Fine Structure of Psychological Time,” *Annals of the New York Academy of Sciences*, vol. 138, no. 2 Interdiscipli, 1967, pp. 623-631.
199. B. Van Bockstaele, et al., “On the costs and benefits of directing attention towards or away from threat-related stimuli: A classical conditioning experiment,” *Behav Res Ther*, vol. 48, 2010, pp. 692-697.
200. S. Martens and B. Wyble, “The attentional blink: past, present, and future of a blind spot in perceptual awareness,” *Neuroscience & Biobehavioral Reviews*, vol. 34, no. 6, 2010, pp. 947-957.
201. A.K. Anderson, “Affective influences on the attentional dynamics supporting awareness,” *Journal of Experimental Psychology: General*, vol. 134, no. 2, 2005, pp. 258-258.
202. P.E. Dux and R. Marois, “The attentional blink: A review of data and theory,” *Attention, Perception, & Psychophysics*, vol. 71, no. 8, 2009, pp. 1683-1700.
203. M.C. Potter, et al., “Picture detection in rapid serial visual presentation: Features or identity?,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 36, no. 6, 2010, pp. 1486-1494; DOI 10.1037/a0018730.
204. M.C. Potter, et al., “Whole report versus partial report in RSVP sentences,” *Journal of memory and language*, vol. 58, no. 4, 2008, pp. 907-915.
205. N.A. Taatgen, et al., “Too much control can hurt: A threaded cognition model of the attentional blink,” *Cogn Psychol*, 2009; DOI S0010-0285(09)00003-6 [pii]
10.1016/j.cogpsych.2008.12.002.
206. W. Duch, et al., “Cognitive Architectures: Where Do We Go From Here?,” *Artificial general Intelligence, 2008: proceedings of the First AGI Conference*, P. Wang, et al., eds., 2008, pp. 122-137.
207. A.V. Samsonovich, “Toward a Unified Catalog of Implemented Cognitive Architectures,” *Proceeding of the 2010 Conference on Biologically Inspired Cognitive Architectures*, A. V. Samsonovich, et al., eds., IOS Press, 2010, pp. 195-244.
208. R.J. Oentaryo and M. Pasquier, “Towards a novel integrated neuro-cognitive architecture (INCA),” 2008, pp. 1902-1909.
209. J.R. Anderson and C. Lebiere, *The atomic components of thought*, Erlbaum, 1998.
210. S. Hélie and R. Sun, “Incubation, insight, and creative problem solving: A unified theory and a connectionist model,” *Psychological Review*, vol. 117, no. 3, 2010, pp. 994-1024; DOI 10.1037/a0019532.
211. R. Sun, et al., “The interaction of the explicit and the implicit in skill learning: A dual-process approach,” *Psychological Review*, vol. 112, no. 1, 2005, pp. 159-192.
212. D.E. Meyer and D.E. Kieras, “A computational theory of executive control processes and human multiple-task performance: Part 1. Basic Mechanisms,” *Psychological Review*, vol. 104, no. 3-65, 1997.

213. P. Langley, et al., “A design for the ICARUS architecture,” *ACM SIGART Bulletin*, vol. 2, 1991, pp. 104–109.
214. L.A. Coward and R. Sun, “Criteria for an effective theory of consciousness and some preliminary attempts,” *Consciousness and Cognition*, vol. 13, no. 2, 2004, pp. 268-301.
215. T. Madl, et al., “Spatial Working Memory in the LIDA Cognitive Architecture,” *Proc. International Conference on Cognitive Modelling*, in review.
216. U. Ramamurthy and S. Franklin, “Memory Systems for Cognitive Agents,” *Proc. Human Memory for Artificial Agents Symposium at the Artificial Intelligence and Simulation of Behavior Convention (AISB'11)*, 2011, pp. 35-40.
217. A.M. Nuxoll and J.E. Laird, “Extending cognitive architecture with episodic memory,” AAAI Press, 2007, pp. 1560-1565.
218. A. Negatu, et al., “A non-routine problem solving mechanism for a general cognitive agent architecture,” *Problem Solving: Techniques, Steps, and Processes*, F. Columbus, ed., Nova Science Publishers, 2011 in press.
219. G. Baldassare, et al., “The IM-CLeVeR project: Intrinsically motivated cumulative learning versatile robots,” 2009, pp. 189-190.
220. A. Baddeley, “Working memory,” *Science*, vol. 255, no. 5044, 1992, pp. 556-559.
221. Sidney K. D'Mello, et al., “A Cognitive Science Based Machine Learning Architecture,” *Proc. AAAI 2006 Spring Symposium Series Sponsor: American Association for Artificial Intelligence.*, 2006.
222. J. Ball, et al., “Toward a large-scale model of language comprehension in ACT-R 6,” *Proceedings of the 8th international conference on cognitive modeling*, 2007.
223. D. Lonsdale, et al., “Resolving a syntactic ambiguity type with semantic memory,” *Book Resolving a syntactic ambiguity type with semantic memory*, Series Resolving a syntactic ambiguity type with semantic memory, ed., 2011, pp. 288-289.
224. R. Lewis, “An architecturally-based theory of human sentence comprehension,” *Book An architecturally-based theory of human sentence comprehension*, Series An architecturally-based theory of human sentence comprehension, ed., 1993, pp.
225. P.S. Rosenbloom, “Rethinking cognitive architecture via graphical models,” *Cogn Syst Res*, vol. 12, no. 2, 2011, pp. 198-209; DOI http://dx.doi.org/10.1016/j.cogsys.2010.07.006.
226. R. Sun, et al., “Modeling meta-cognition in a cognitive architecture,” *Cogn Syst Res*, vol. 7, no. 4, 2006, pp. 327-338.
227. Z. Zhang, et al., “Metacognition in Software Agents using Classifier Systems,” *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, MIT Press, 1998, pp. 83–88.
228. Lee D. Erman, et al., “The Hearsay-II Speech-Understanding System: Integrating Knowledge to Resolve Uncertainty,” *Computing Surveys*, vol. 12, 1980, pp. 213–253.
229. J.L. McClelland, “The programmable blackboard model of reading,” *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 2, 1986, pp. 122-169.
230. M. de Kamps and F. van der Velde, “Neural blackboard architectures: the realization of compositionality an∗d systematicity in neural networks,” *J. Neural Eng.*, vol. 3, 2006, pp. R1–R12; DOI 10.1088/1741-2560/3/1/R01.
231. H.P. Nii, “The Blackboard Model of Problem Solving and the Evolution of Blackboard Architectures,” *The AI Magazine*, vol. Summer 1986, 1986, pp. 38–53.
232. S. Dehaene, et al., “A neuronal network model linking subjective reports and objective physiological data during conscious perception,” *Proc. Natl. Acad. Sci. USA*, vol. 1001, 2003, pp. 8520–8525.
233. R. Wallace, *Consciousness: A Mathematical Treatment of the Global Neuronal Workspace Model*, Springer, 2005.
234. D. Connor and M. Shanahan, “A computational model of a global neuronal workspace with stochastic connections,” *Neural Networks*, vol. 23, no. 10, 2010, pp. 1139-1154; DOI 10.1016/j.neunet.2010.07.005.
235. J. Snaider, et al., “Time production and representation in a conceptual and computational cognitive model,” *Cogn Syst Res*, vol. 13, no. 1, 2012, pp. 59-71; DOI 10.1016/j.cogsys.2010.10.004.
236. S. Franklin, “Automating Human Information Agents,” *Practical Applications of Intelligent Agents*, Z. Chen and L. C. Jain, eds., Springer-Verlag, 2001, pp. 27–58
237. M. Beetz, et al., “CRAM - A Cognitive Robot Abstract Machine for Everyday Manipulation in Human Environments,” *Ieee Int C Int Robot*, 2010, pp. 1012-1017.

**Stan Franklin**, a mathematician turned computer scientist turning cognitive scientist, holds the W. Harry Feinstone Interdisciplinary Research Professorship at the University of Memphis, was a founder of its Computer Science Department and of its Institute for Intelligent Systems. Wanting to know how minds work — human minds, animal minds and, particularly, artificial minds — motivates his research on cognitive modeling. His MS and PhD degrees are from UCLA, his undergraduate degree from the University of Memphis. He has authored or co-authored some hundred and fifty academic papers that have been cited over 7000 times. He serves on the editorial boards of seven international professional journals including *Topics in Cognitive Science*, the *Journal of Artificial General Intelligence*, *Biologically Inspired Cognitive Architectures*, the *Journal of Mind Theory*, and the *International Journal of Machine Consciousness.*

**Tamas Madl** is a postgraduate student at the University of Manchester, a postgraduate researcher at the Austrian Institute for Artificial Intelligence, and an external member of the Cognitive Computing Research Group at the University of Memphis. He has earned a BSc degree in computer science from the University of Central Lancashire in 2009, and an MSc degree in cognitive science from the University of Vienna in 2011.

His current research interests include spatial memory and navigation in cognitive architectures, models of functional consciousness and of the perception-action cycle, and Bayesian approaches to cognitive modeling. He has previously worked on adjusting the LIDA cognitive architecture’s timing parameters, and has implemented several LIDA-based cognitive software agents reproducing psychological experiments.

**Sidney D’Mello** is an Assistant Professor in the departments of Computer Science and Psychology at the University of Notre Dame. His interests include affective computing, artificial intelligence, human-computer interaction, speech recognition, and natural language understanding. He has published over 150 journal papers, book chapters, and conference proceedings. He is an associate editor for *IEEE Transactions on Affective Computing* and an advisory editor for the *Journal of Educational Psychology*. D’Mello received his PhD. in Computer Science from the University of Memphis in 2009.

**Javier Snaider** earned an Electronics Engineering degree, with honors diploma, from the Universidad de Buenos Aires, Argentina in 1992, and Computer Science PhD degree from the University of Memphis in 2012. He is also member of the Cognitive Computing Research Group.

He has worked as a computer software consultant, particularly for Java projects. He is a certified Java instructor and as such, he has taught advanced Java courses for enterprises such as Sun and Oracle. He also taught at the University of Buenos Aires and at the Universidad Tecnologica Nacional (Argentina).

His primary research interest is in cognitive architectures, their internal representations, memory and learning algorithms, and their cognitive processes. He developed a new Integer Sparse Distributed Memory, which replaces the original binary vectors of SDM with modular integer vectors. He also created the Modular Composite Representation, a high-dimensional vector representation that is able to represent complex structures using single vectors.

His work with the LIDA cognitive architecture is related to time perception and production, its memory system, its main representation model, and the design and implementation of the LIDA computational framework, a highly customizable Java implementation of the LIDA architecture. Currently, he is working at Google Inc. in Mountain View CA. He is developing new API services for several platforms, including web and Android mobile operating system.