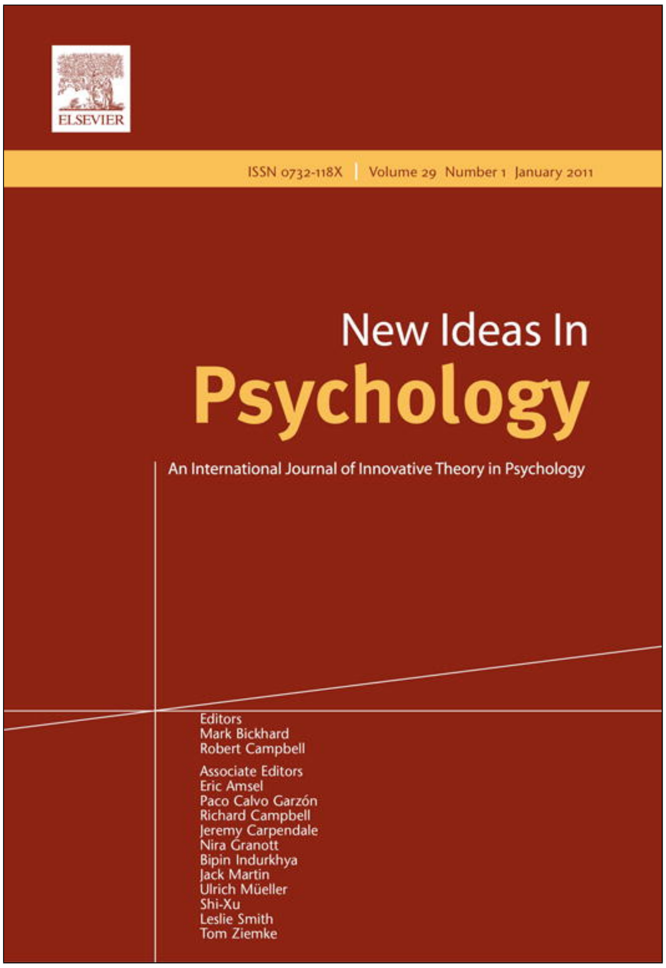

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

New Ideas in Psychology 30 (2012) 227–240

Contents lists available at SciVerse ScienceDirect

New Ideas in Psychology

journal homepage: www.elsevier.com/locate/newideapsych

Memory systems within a cognitive architecture

Ron Sun*

Cognitive Science Department, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

Keywords:
Memory
Cognitive architecture
CLARION

A B S T R A C T

This article addresses the division of memory systems in relation to an overall cognitive architecture. As understanding the architecture is essential to understanding the mind, developing computational cognitive architectures is an important enterprise in computational psychology (computational cognitive modeling). The article proposes a set of hypotheses concerning memory systems from the standpoint of a cognitive architecture, in particular, the four-way division of memory (including explicit and implicit procedural memory and explicit and implicit declarative memory). It then discusses in detail how these hypotheses may be validated through examining qualitatively the literature on memory. A quick review follows of computational simulations of a variety of quantitative data (which are not limited to narrowly conceived "memory tasks"). Results of accounting for both qualitative and quantitative data point to the promise of this approach.

1. Introduction

Human memory systems have evolved in the process of coping with the world (like other parts of the mind; Clark, 1997; Dreyfus, 1992; Klein, Cosmides, Tooby, & Chance, 2002; Sun, 2002; Vicente & Wang, 1998). Thus, memory systems have resulted from the informational needs of cognitive agents in their coping with the world; this point has been extensively argued before (e.g., Hirschfield & Gelman, 1994; Klein et al., 2002; Sun, 2002; Vicente & Wang, 1998). Consequently, memory systems serve the purpose of supplying useful information relevant to the current activities of the cognitive agent in a pertinent and timely manner. A characteristic of memory is more likely to exist if it contributes to the essential activities of an agent, and is less likely if it does not. Moreover, different activities of agents may require different memory properties and therefore possibly different memory systems to provide corresponding properties. In this regard, a non-memory-centered way of looking at memory is needed in order to better understand it from the standpoint of cognitive agents' essential activities (Dreyfus, 1992; Klein et al., 2002; Sun, 2002). That is, we need a broader perspective (e.g., from the standpoint of a cognitive architecture; Newell, 1990).

As some researchers have argued on the basis of empirical data (e.g., Klein et al., 2002; Schneider, 1993), multiple memory systems have co-evolved to function in a complementary (and often inter-locking) way. For example, when a generalization is retrieved from semantic memory, exceptions may be retrieved from episodic memory to provide complementary information. For another example, dissociations among memory systems may exist under some circumstances, so that different memory systems may serve different purposes, but dissociations may not exist under some other circumstances, so that multiple memory systems may be brought together to bear on one task. Dissociation may also lead to synergy when memory systems are properly integrated (e.g., Sun, Slusarz, & Terry, 2005).

Each memory system solves some specific problems that cognitive agents encounter in the course of evolution. Investigating the problems faced by each memory system may help to achieve a better understanding and delineate the design features of each memory system. In other words, we may achieve a clearer mapping of functions to systems

* Tel.: +1 518 276 3409; fax: +1 518 276 8268.
E-mail address: rsun@rpi.edu.

0732-118X/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.newideapsych.2011.11.003

in this way (specifically with regard to memory). As pointed out by Klein et al. (2002), "agnostically cataloging arbitrary samples out of the inexhaustible set of everything a memory system is capable of doing is not likely to lead to knowledge of its function". A system of any kind may be capable of doing an endless series of things that it was not designed to do. Therefore, we cannot simply study memory by exhaustively testing what it is capable of. That is, we need to take a look at memory systems from a non-memory-centered, broader perspective. On that basis, a cognitive architecture needs to be specified so as to structurally situate a variety of ecologically realistic or plausible memory systems. Note that memory systems may, alternately, be viewed either as functional modules or as physiological modules (for more discussions of this and other distinctions, see Fodor, 1983; Sun, 2009). However, my main focus here is on functional modules.

As we know, a cognitive architecture is the essential structures and processes of cognition in the form of a broadly-scoped, (more or less) domain-generic computational model, used for a broad, multiple-level, multiple-domain analysis of behavior. A cognitive architecture provides a concrete framework for more detailed modeling of cognitive phenomena, through specifying essential structures, divisions of modules, relations among modules, and a variety of other aspects (Sun, 2004). Cognitive architectures are believed to be essential in advancing understanding of the mind (Anderson & Lebiere, 1998; Newell, 1990; Sun, 2002, 2004). Therefore, developing cognitive architectures has been an important enterprise in cognitive science. We can benefit a great deal from utilizing cognitive architectures for the following reasons. First of all, a cognitive architecture provides a functional division of modules. Second, cognitive architectures are functionally comprehensive (relatively speaking and to a certain extent). Third, cognitive architectures are generic (domain-general) to a large extent. Correspondingly, the analysis of memory systems through cognitive architectures can be performed at the computational, or functional, level, on the basis of empirical data. Based on the data, hypotheses concerning memory systems can be tested through simulating a wide range of data, not just from traditional "memory tasks", but also from other psychological tasks that rely on memory in various ways, perhaps better reflecting cognitive agents' everyday interaction with the world.

Many issues need clarification in this process. For example, it is far from clear what essential subsystems of memory are and how memory should be divided up functionally. Some other issues concern essential dichotomies in cognition. Various cognitive dichotomies have been proposed before: implicit versus explicit, procedural versus declarative, automatic versus controlled, and so on. What are the essential dichotomies? How should we analyze them in a process-based (computational, mechanistic) way, and thereby develop cognitive architectures on that basis? Many other issues exist as well. See, for example, Anderson and Lebiere (1998), Newell (1990), and Sun (2004) for some of these other issues. In this article, I will focus mainly on functional-structural issues (that is, issues such as essential cognitive dichotomies, modularity, memory systems/modules, and so on).

In this work, I try to shed light on some of these issues with the help of computational modeling. As pointed out by Hintzman (1990), "the common strategy of trying to reason backward from behavior to underlying processes (analysis) has drawbacks that become painfully apparent to those who work with simulation models (synthesis). To have one's hunches about how a simple combination of processes will behave repeatedly dashed by one's own computer program is a humbling experience that no experimental psychologist should miss" (p. 111). "A simple working system that displays some properties of human memory may suggest other properties that no one ever thought of testing for, may offer novel explanations for known phenomena, and may provide insight into which modifications that the next generation of models should include" (p. 111). The importance of computational perspectives should not be under-estimated.

In the remainder of this article, I will first, in Section 2, present an overview of a framework—a synthesis in the form of a cognitive architecture and the memory systems within. Section 3 details the empirical support for this framework of memory systems, synthesizing empirical research. Section 4 provides a general discussion, and concludes this article.

2. Overview of memory systems in CLARION

CLARION is an attempt in the afore-mentioned direction. It specifies various functional modules ("subsystems" and modules within), with their corresponding memory systems, in an overall architectural framework (Sun, 2002, 2003), in accordance with an ecological-functional perspective (Sun, 2002, 2004).[1] Although CLARION has been described in bits and pieces in psychological and cognitive science journals before, a synthesis as in the present work is necessary.

CLARION, as a computational model, allows more precise, mechanistic (i.e., computational) specifications of memory systems, which are (at least) algorithmically clearer, more consistent, and working. Let us first review some general ideas regarding CLARION.

2.1. Overall structure

Overall, CLARION is an integrative cognitive architecture consisting of a number of distinct "subsystems", with a dual representational memory structure in each subsystem (implicit versus explicit representations; see Cleeremans, Destrebecqz, & Boyer, 1998; Reber, 1989; Seger, 1994; Sun et al., 2005). CLARION is intended for capturing all the

[1] As variously discussed in Sun (2002, 2004), the major tenets of the ecological-functional perspective may include: (a) taking the role of function as a starting point, (b) considering evolution seriously, (c) paying particular attention to the ecological niche (at present or evolutionarily), (d) focusing on routine everyday activities that are most representative of the ecological niche, (e) basing theories on the notion that cognitive characteristics are functional, that is, useful in some way for routine everyday activities within the ecological niche, (f) basing theories on the notion that cognitive characteristics are often selected based on cost-benefit tradeoffs, at present or evolutionarily.

essential cognitive processes within an individual cognitive agent in its routine, everyday activities (Dreyfus, 1992; Heidegger, 1927), in accordance with the ecological-functional perspective (Sun, 2002, 2004). Accordingly, its subsystems include the action-centered subsystem (the ACS), the non-action-centered subsystem (the NACS), the motivational subsystem (the MS), and the metacognitive subsystem (the MCS). The role of the ACS is to control actions (based on procedural knowledge), regardless of whether the actions are for external physical movements or internal mental operations. The role of the NACS is to maintain general (i.e., declarative) knowledge (ultimately in the service of action decision making by the ACS). The role of the MS is to provide underlying motivations for perception, action, and cognition. The role of the MCS is to monitor, direct, and modify the operations of the other subsystems, for the sake of better performance. Each of these subsystems, in an ecological-functional sense, serves a unique function, and together they form a functioning cognitive architecture. See Fig. 1 (see Sun, 2002, 2003 for computational details and demonstrations).

The more information a cognitive agent encodes, the more difficult it is to devise a system that can deliver the right information at the right moment (and do so without overwhelming it with less relevant or irrelevant materials; Klein et al., 2002). Modularization of memory such as the division above (see more discussions below) is therefore of importance (see Coward & Sun, 2004 regarding issues of modularization). Similarly, different informational needs of an agent may require different memory processes and different memory characteristics, and therefore likely different memory modules/systems. Memory systems and characteristics that have little or no impact on an agent's actions and activities likely will not be evolutionarily selected. Therefore, the issues of memory modules/systems should be approached with the "adaptive" functions of the different memory systems in mind within the overall cognitive architecture.

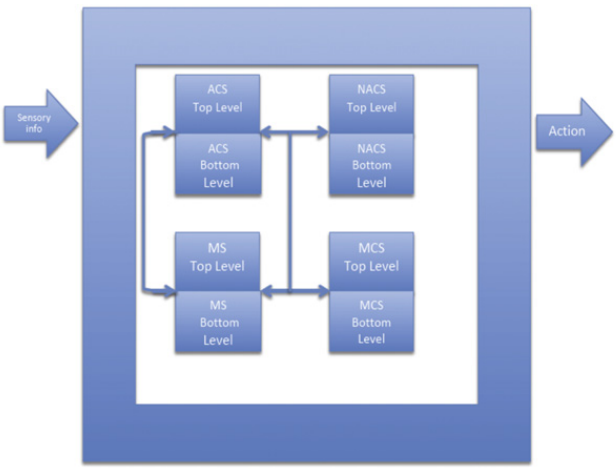

Fig. 1. The subsystems of the CLARION cognitive architecture. The major information flows are shown with arrows. See Sun (2003) for full technical details.

2.2. Two levels

Each of these subsystems consists of two "levels" of representation (i.e., two memory modules/systems), which together constitute a dual representational (dual memory) structure. See Fig. 1. Generally, in each subsystem, the top "level" encodes explicit knowledge (explicit memory) and the bottom "level" encodes implicit knowledge (implicit memory). This distinction has been argued for amply before (see Cleeremans et al., 1998; Reber, 1989; Seger, 1994; Sun, 2002; Sun et al., 2005), and we will further justify it later (see the later section on accounting for memory phenomena qualitatively).

The relatively inaccessible nature of implicit knowledge in implicit memory may be captured by subsymbolic, distributed representation provided, for example, by a backpropagation network (Rumelhart, McClelland, & the PDP Research Group, 1986). This is because distributed representational units in the hidden layer(s) of a backpropagation network are capable of accomplishing computations but are subsymbolic and generally not individually meaningful (Rumelhart et al., 1986; Sun, 1994). This characteristic of distributed representation, which renders the representational form less accessible, accords well with the relative inaccessibility of implicit knowledge (Cleeremans et al., 1998). In contrast, explicit knowledge in explicit memory may be captured (in computational modeling) by symbolic or localist representation, in which each unit is more easily interpretable and has a clearer conceptual meaning. This characteristic of symbolic or localist representation captures the characteristic of explicit knowledge being more accessible and more manipulable (Sun, 1994).

Accessibility here refers to the direct and immediate availability of mental content for the major operations that are responsible for, or concomitant with, consciousness, such as introspection, forming higher-order thoughts, and verbal reporting.

The dichotomous difference in the representations of the two different types of memory leads naturally to a two-level structuring within each subsystem, whereby each level (module) uses one kind of representation and captures one corresponding type of memory process (implicit or explicit).

In relation to the ecological-functional perspective, there have been some indications that one level (the explicit level) is evolutionarily newer than the other (the implicit level) (see, e.g., LeDoux, 1996). Beside this evolutionary difference, the juxtaposition of the two levels is also functional. It is functional (and thus evolutionarily advantageous), especially because, as I have extensively argued in prior work, the interaction between the two levels may lead to synergy in the form of better, more accurate, and/or faster performance in a variety of circumstances (see Sun, 2002; Sun, Merrill, & Peterson, 2001, Sun et al., 2005 for details). Therefore, such a division between the two levels may conceivably be favored by natural selection.

2.3. Multiple memory systems

In the main, within the action-centered and the non-action-centered subsystem, a four-way partitioning of the major memory systems (modules) is as follows: implicit procedural (action-centered) memory (in the ACS), explicit procedural (action-centered) memory (in the ACS), implicit declarative (non-action-centered) memory (in the NACS), and explicit declarative (non-action-centered) memory (in the NACS). (For the time being, I ignore the memories within the motivational and the metacognitive subsystem, as well as working memory and so on.)

However, declarative (non-action-centered) memory may be further partitioned into semantic memory and episodic memory (Tulving, 1985; with the implicit-explicit distinction in each as will be discussed later). Semantic memory is used for storing general (non-action-centered) knowledge that is not tied to specific experiences, while episodic memory is for storing experience-specific information. The distinction between episodic memory and semantic memory in CLARION is justified, in an ecological-functional sense, based on the need to separate knowledge that is concerned with specific experiences and knowledge that is more generic and thus not experience-specific (see, e.g., Klein et al., 2002 for justifications along this line), although some memory researchers did not make such a distinction (e.g., Hintzman, 1986).

In addition to long-term memory systems (explicit or implicit) as enumerated above, note that there is also a working memory in CLARION, which is for storing information temporarily for the purpose of facilitating action decision making (roughly along the line of Baddeley, 1986). In particular, sensory information store, a part of the working memory in CLARION, is for storing sensory input information temporarily. Goal structures, a special case of the working memory in CLARION, are for storing goal information specifically (e.g., Anderson & Lebiere, 1998). This module is justifiable in an ecological-functional sense from the description above (Sun, 2002).

Thus, Table 1 presents the taxonomy of memory modules in CLARION. Fig. 2 presents a diagrammatic sketch of the four major long-term memory systems interacting with each other (within the ACS and the NACS of CLARION), whereby declarative memory is further partitioned into semantic and episodic memory. See Sun (2003) for technical details. Note that modularity implies certain information encapsulation:

Table 1
A proposal concerning the taxonomy of memory according to CLARION.

Long-term memory:			
Explicit:	Procedural memory	Semantic memory	Episodic memory
Implicit:	Procedural memory	Semantic memory	Episodic memory
Working memory:			
Sensory input			
Goal structure			
Other temporary info			

Memory type	Representational entity	Type of representation
Implicit memories	Connectionist networks	Distributed representation
Explicit memories	Rules and chunks	Symbolic/localist representation

in CLARION, modules only operate on certain kinds of inputs, and they operate relatively independently. Information coming into and going out of a module is relatively limited. Dissociations may thus follow as a result.

This view of memory systems is certainly not parsimonious. We need to be aware of the limitations of any parsimony argument. The human mind/brain has evolved incrementally over a long period of time. New systems are often the results of mutations of old systems (not of completely new design), and often function on top of (or along with) existing systems. In addition, different mutations of an old system may lead to several different (but co-existing) new systems. The multiple systems that result from mutations may exhibit significant differences, but may overlap in functionality and be complementary to each other (Klein et al., 2002; McClelland, McNaughton, & O'Reilly, 1995; Sun, 2002). Schneider (1993) made a similar argument in relation to visual systems, and pointed out that "it is likely that the biology of memory is analogously composed of multiple overlapping memory systems that are optimized for different classes of information storage" (p. 184).

Note that this division of memory systems is very different from some existing views that have been well entrenched. For example, ACT-R has been adhering to a procedural-declarative dichotomy, and it is often claimed that procedural memory in ACT-R is implicit and in declarative memory explicit and implicit information is tightly coupled item by item (more on ACT-R later; see Anderson & Lebiere, 1998). CLARION is also distinguished from other existing frameworks (such as SOAR; Rosenbloom, Laird, & Newell, 1993). See also Logan (2002) and Ratcliff and McKoon (1998).

2.4. Multiple learning processes

Given the partitioning of memory systems above, learning in these memory systems need to be mentioned (also in accordance with the ecological-functional perspective). First, there is the learning of implicit knowledge at the bottom level of a subsystem (i.e., in implicit memory). One way of implementing a mapping function to capture implicit knowledge in implicit memories is to use a multi-layer neural network (e.g., a three-layer back-propagation network). Adjusting parameters of this function to change input/output mappings (that is, to learn implicit knowledge) may be carried out in ways consistent with the nature of distributed representation, through trial-and-error interaction with the world (in keeping with the ecological realism consideration; Sun, 2002). For instance, reinforcement learning (as developed in machine learning) may be appropriate for capturing human learning of implicit action-centered knowledge, especially Q-learning (Watkins, 1989), implemented using back-propagation networks (Sun et al., 2001). Such implicit learning may be justified from the ecological-functional perspective: For instance, Cleeremans (1997) argued at length that implicit learning could not be captured by symbolic models but neural networks; Sun (1999, 2002) made similar arguments.

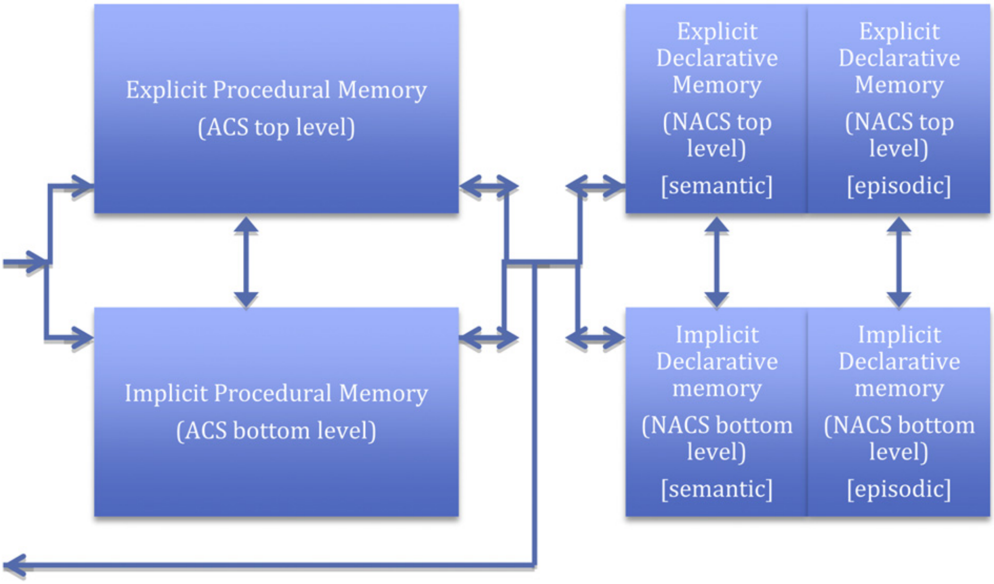

Fig. 2. The essential memory modules of the CLARION cognitive architecture. The leftmost lines show the input information to and output actions from the ACS. The lines between the modules show the information flows (the working memory, goal structure, and sensory information store are used to facilitate the flows but are omitted above for the sake of simplicity). Note that the memory modules are relatively encapsulated, and information flows, as indicated by arrows, are limited between modules. See Sun (2003) for full technical details.

Explicit knowledge at the top level of a subsystem (i.e., in explicit memory) can also be learned in a variety of ways (in accordance with localist/symbolic representation used there). Because of its representational and other characteristics, one-shot learning (e.g., based on hypothesis testing) may be appropriate during the interaction with the world (see, e.g., Bruner, Goodnow, & Austin, 1956; Sun et al., 2001). With such learning, a cognitive agent dynamically acquires explicit representations and modifies them as needed (in keeping with the ecological realism consideration; Sun, 2002).

The implicit knowledge already acquired in the bottom level (i.e., in implicit memory) may be utilized in learning explicit knowledge at the top level (in explicit memory), through "bottom-up learning" (Sun et al., 2001). That is, implicit information accumulated through interacting with the world is used for extracting and then refining explicit knowledge. This is a kind of "rational reconstruction" of implicit knowledge at the explicit level. Other types of learning of explicit knowledge in explicit memory are also possible, such as explicit hypothesis testing without the help of the bottom level. Conversely, once explicit knowledge is established at the top level (in explicit memory), it may be assimilated into the bottom level (implicit memory). This often occurs during the novice-to-expert transition in instructed learning settings (Anderson & Lebiere, 1998; Dreyfus & Dreyfus, 1987). The assimilation process, termed "top-down learning" (as opposed to "bottom-up learning"), may be carried out in a variety of ways (see Sun, 2002, 2003 for details).

Memory systems in CLARION, as in humans, are there to serve the need of coping with everyday activities: They actively "participate" in the activities of the agent through meeting the informational needs of the agent arising from such activities. Learning in these memory systems as described above should be seen in this light (in accordance with the ecological-functional perspective).

2.5. Major hypotheses

To summarize the major hypotheses regarding memory systems from the CLARION cognitive architecture:

- **Hypothesis I**: the distinction between implicit and explicit memory;
- **Hypothesis II**: the distinction between action-centered (procedural) and non-action-centered (declarative) memory, which is hypothesized to be orthogonal to the first distinction above;
- **Hypothesis III**: the distinction between semantic and episodic memory (within declarative memory), which is also hypothesized to be orthogonal to the first distinction above.

Below, evidence related to these hypotheses will be evaluated. (Details of these hypothesized memory systems have been described in bits and pieces in Hélie & Sun, 2010; Sun, 2003; Sun et al., 2005.)

With a computational perspective, the present work introduces a number of new ideas, including the four-way division of memory (including explicit and implicit procedural memory and explicit and implicit declarative memory), as well as other auxiliary memory modules. These memory modules are significantly different from other existing cognitive architectures (such as ACT-R, as detailed later). Another distinction unique to CLARION is bottom-up versus top-down learning through the interaction between implicit and explicit memory modules (see, e.g., Sun, 2002; Sun et al., 2001 for details; more later).

3. Accounting for memory phenomena qualitatively

3.1. Validating the hypotheses

The three major hypotheses of CLARION are examined below in relation to available empirical evidence.

3.1.1. Evidence in support of hypothesis I

We first examine the distinction between implicit and explicit memory. The theoretical distinction between implicit and explicit processes, as well as its ecological-functional significance, has been variously argued in the past in many empirically-based theories and models, for example, in Reber (1989), Seger (1994), as well as in Sun (1994, 2002).

First of all, the distinction of implicit and explicit processes has been empirically demonstrated in the implicit memory literature (Roediger, 1990; Schacter, 1987). For example, the work on amnesics in the 60s and 70s showed that they might have intact implicit memory while their explicit memory was severely impaired. Warrington and Weiskrantz (1970) demonstrated that amnesics' memory using implicit measures was as good as normals while their memory using explicit measures was worse than normals. The explicit measure they used included free recall and recognition, while the implicit measures they used included word fragment naming and word completion. It might be argued that their implicit measures reflected unconscious (implicit) processes because amnesic patients were usually unaware that they knew the materials. The results, while surprising in the 70s, have been replicated many times since then.

Jacoby (e.g., 1983) demonstrated that implicit and explicit measures might be dissociated among normal subjects as well (see also Dunn & Kirsner, 1988). Three study conditions were used: generation of a word from a context, reading aloud a word in a meaningful context, and reading aloud a word out of context. The dependent measures were recognition (from a list of words) and perceptual identification (from fast presentations of words). Their results showed that using the explicit measure (recognition), generated words were remembered the best and words read out of context were remembered the least. However, using the implicit measure (perceptual identification), the exactly opposite pattern was found. Other similar cross-over dissociations were also found from other manipulations (Roediger, 1990). Toth, Reingold, and Jacoby (1994) further devised the inclusion-exclusion procedure, which provided even stronger indications of the dissociation of implicit memory from explicit memory.

The distinction of implicit and explicit processes has also been empirically demonstrated in the implicit learning literature (e.g., Berry & Broadbent, 1988; Cleeremans et al., 1998; Reber, 1989; Seger, 1994). There have been three common tasks used in implicit learning research. For example, serial reaction time tasks (Willingham, Nissen, & Bullemer, 1989) probe learning of a repeating sequence. It was found that there was a significant reduction in response time to repeating sequences (relative to random sequences). However, participants might not be able to explicitly report the repeating sequence, and were often unaware that a repeating sequence was involved (see, e.g., Lewicki et al., 1987). Similarly, process control tasks (Berry & Broadbent, 1988) examine learning of a relation between the input and the output variables of a controllable system, through interacting with the system dynamically. Although they often did not recognize the underlying relations explicitly, participants reached a certain level of performance in these tasks. Likewise, in artificial grammar learning tasks (Reber, 1989), participants were presented with strings of letters that were generated in accordance with a finite state grammar. After memorization, participants showed an ability to distinguish new strings that conformed to the artificial grammar used to generate the initial strings from those that did not, although participants might not be explicitly aware of the underlying grammar (except for some fragmentary knowledge possibly). In all, these tasks share the characteristic of implicit performance (at least to a significant extent).

There are many other tasks that are similar in this regard, such as various concept learning, reasoning, automatization, and instrumental conditioning tasks (see Sun, 2002 for a review; see also Evans & Frankish, 2009). Together, they demonstrated the distinction between implicit and explicit processes. Although some researchers have disputed the existence of implicit processes based on the imperfection and incompleteness of tests for explicit knowledge (e.g., Shanks & St.John, 1994), there is an overwhelming amount of evidence in support of the distinction between implicit and explicit processes (Sun et al., 2005).

Several other theoretical distinctions made by other researchers capture some similar differences between two different types of cognitive processes. Dreyfus and Dreyfus (1987) proposed the distinction of analytical and intuitive thinking, and believed that the transition from the former to the latter was essential to the development of complex cognitive skills, on the basis of phenomenological analysis of different stages of learning to play chess. Similar distinctions have been proposed by other researchers based on different empirical or theoretical considerations. See, for example, Evans and Frankish (2009) for a collection of such theories. Taken together, the distinction between explicit and implicit processes may be supported in many ways (although details of some of these proposals might be different or even contradictory to each other in some way).

Now the question is whether these different types of memory reside in separate memory systems (modules) or not. There have been debates (see, e.g., Roediger, 1990), and differing views exist. Squire (1987) proposed that memory be divided into declarative and procedural memory, with the former further divided into episodic (working) and semantic (reference) memory and the latter into skills, priming, classical conditioning, and other memory stores. According to Squire (1987), declarative memory is explicit while procedural memory is implicit. Tulving and Schacter (1990) incorporated some features of the one-system view (e.g., the transfer-appropriate-processing view), while preserving the separation of explicit and implicit memory. They proposed that there should be multiple priming systems in the implicit memory, so that dissociations among different implicit measures could be accounted for (e.g., the priming system should be divided into the perceptual and the structural description system). This proposal addressed some of the objections raised by the proponents of the one-system view. Sun (2002) and Sun et al. (2005) provided some theoretical interpretations of existing learning data (related to process control, serial reaction time, and other tasks), based on a multiple memory systems view.

It is worth noting that in social psychology, there have been a number of dual-process models that are in fact roughly based on the co-existence of implicit and explicit memory systems (see, for example, Chaiken & Trope, 1999; see also Evans & Frankish, 2009).

I will show later that this division is indeed functional (in accordance with the ecological-functional perspective; in the Sections 3.2.1 and 3.2.2). The idea is that the separation of the two types of information enables the application of each type of information separately as appropriate for different types of situations. For example, highly complex situations may be better handled by using information in implicit memory, while information in explicit memory may be better for more clear-cut situations because such information may be used in a more precise way (see Sun, 2002; Sun & Mathews, 2005). Furthermore, the interaction of the two separate types of information may lead to overall better performance, that is, synergy between the two types, under proper circumstances (as demonstrated in, e.g., Mathews et al., 1989; Sun et al., 2001, 2005). So the separation and the interaction of these two types of memory may be cognitively advantageous (Sun, 2002).

3.1.2. Evidence in support of hypothesis II

I now turn to the distinction between procedural and declarative memory (action-centered and non-action-centered memory in the ACS and the NACS, respectively) in CLARION and its orthogonality with the implicit-explicit distinction (the latter might be more controversial).

First, the distinction between procedural and declarative memory has been proposed by Anderson (1983), Squire (1987), and others (although some details vary across these proposals). Procedural memory contains knowledge that is specifically concerned with actions (and action sequences) in various circumstances, that is, how to do things. Declarative memory contains knowledge that is not specifically concerned with actions (or action sequences), but more about objects, events, and so on in generic terms (i.e., the "what", not the "how"). The major factor that distinguishes procedural and declarative memory seems to be the action-centeredness or the lack thereof (or in other words, the procedural versus non-procedural nature of knowledge).

Evidence in support of this distinction includes voluminous studies of skill acquisition in both high- and low-level skill domains (e.g., Ackerman & Kanfer, 2004; Anderson, 1983; Anderson & Lebiere, 1998; Kanfer & Ackerman, 1989; Sun et al., 2001; and so on; see also Proctor & Dutta, 1995). These studies include both experimental work on human subjects, as well as modeling/simulation and other work aimed at theoretical interpretations. They showed that making this distinction provides useful insight in interpreting a range of data and phenomena. For instance, Anderson (1983) proposed this distinction to account for changes in performance resulting from extensive practice, based on data from a variety of skill learning studies (ranging from arithmetic to geometric theorem proving). For Anderson, the initial stage of skill development is characterized by the acquisition of declarative knowledge (explicit knowledge concerning a task). During this stage, the learner must explicitly rely on this knowledge in order to successfully perform a task. Through practice, implicit procedures develop that enable the performance of the task without explicit declarative knowledge, and often without concurrent conscious awareness of details involved. By any measure, this distinction is well established (Proctor & Dutta, 1995).

Let us examine the relationship between the procedural-declarative distinction and the implicit-explicit distinction. In Anderson (1983), declarative knowledge in the declarative memory is assumed to be consciously accessible (i.e., explicit): Subjects can manipulate and report on such knowledge. Procedural knowledge in the procedural memory is not: It leads to action without explicit subjective accessibility. Thus, in Anderson (1983), the two dichotomies were merged into one.

In ACT-R as described by Anderson and Lebiere (1998), on the other hand, each individual piece of knowledge, be it procedural or declarative, involves both subsymbolic and symbolic representation. Symbolic representation is used for denoting semantic labels and structural components of each concept, while subsymbolic representation is used for expressing its activation and other numerical factors. One theoretical interpretation is that the symbolic representation is explicit while the subsymbolic representation is implicit (either for declarative knowledge, or for both declarative and procedural knowledge). This view constitutes another take on the relationship between the two dichotomies.

According to the first view above, the difference in action-centeredness (i.e., the procedural versus non-procedural nature) seems the main factor in distinguishing the two types of knowledge in the two different memory systems, while accessibility (i.e., implicitness versus explicitness) is a secondary factor. We believe that this view unnecessarily confounds two aspects: action-centeredness and accessibility, and can be made clearer by separating the two dimensions. There are sufficient reasons to believe that action-centeredness does not necessarily go with implicitness (inaccessibility), as shown by, for example, the experiments of Stanley, Mathews, Buss, and Kotler-Cope (1989), Sun et al. (2001), or Willingham et al. (1989). Likewise, non-action-centeredness does not necessarily go with explicitness (accessibility) either, as shown by conceptual priming and other implicit memory experiments (e.g., Moscovitch & Umilta, 1991; Schacter, 1987) or by the experiments demonstrating implicit statistical information (Hasher & Zacks, 1979; Nisbett & Wilson, 1977). Note that some researchers might group all implicit memory (including semantic, associative, and conceptual priming) under procedural memory (e.g., Squire, 1987), but such views confound the definition of "procedural" and are thus not adopted here. Therefore, I distinguish these two separate dimensions when theorizing on memory.

The alternative ACT-R view that each individual piece of knowledge (either procedural or declarative, or both) involves both implicit and explicit processes is also problematic. Such a framework entails a close-coupling between implicit and explicit processes that is highly questionable. The underlying assumption that every piece of knowledge (either declarative or procedural, or both) has an explicit part contradicts the fact that some knowledge may be completely implicit (see, e.g., Cleeremans et al., 1998; Lewicki et al., 1987). This contradiction raises the question of whether a tight coupling or a more separate organization, for example, by having these two types of knowledge in separate memory modules/systems, makes better sense.

Squire (1987) proposed that memory should be divided into declarative and procedural memory, with the former being further divided into episodic (working) and semantic (reference) memory and the latter into skills, priming, classical conditioning, and other memory stores; declarative memory is explicit while procedural memory is implicit. However, this view would have trouble accounting for implicit declarative memory (which is clearly not procedural; e.g., conceptually driven priming; Roediger, 1990). Explicit procedural memory is also not accounted for in this view (see Sun, 2002; Sun et al., 2005).

As a more natural and more intuitively appealing alternative to those views above, in CLARION, I would propose the separation of the two dichotomies: I would treat them as logically separate from (i.e., orthogonal to) each other (Sun, Zhang, & Mathews, 2009). Arguments in favor of this view can be found in the literature. For example, Willingham (1998) argued based on empirical data that motor skills (a kind of procedural knowledge) consist of both implicit and explicit processes. Rosenbaum, Carlson, and Gilmore (2001) argued based on empirical data that both intellectual skills and perceptual-motor skills are made up of implicit and explicit knowledge. In other words, procedural knowledge (action-centered knowledge), ranging from high-level intellectual skills to perceptual-motor skills, may be divided into implicit and explicit procedural memory.

Similarly, declarative knowledge (non-action-centered knowledge) may also be divided into implicit and explicit memory (Tulving & Schacter, 1990). Functionally, there is no reason to believe that all implicit knowledge is procedural (as advocated by some of the afore-mentioned views). Some of the implicit knowledge may in fact be declarative (non-action-centered). Furthermore, in terms of functional consideration, having separate implicit and explicit declarative memory allows different tasks to be tackled simultaneously in these two separate memory stores (e.g., while thinking explicitly about one task, letting intuition work on another). Sun (1994) and Sun and Zhang (2007) showed that, through dividing declarative memory into explicit and implicit modules, similarity-based reasoning data could be naturally and succinctly accounted for. Furthermore, Hélie and Sun (2010) showed that this division accounted well for creative problem solving (which otherwise would be difficult to account for).

In CLARION, procedural and declarative knowledge reside separately in procedural and declarative memory respectively, which are representationally different. Procedural knowledge (in the ACS) is represented by either action rules (explicit) or action neural networks (implicit), both of which are centered on situation-action mappings. Declarative knowledge (in the NACS), on the other hand, is represented by either associative rules (explicit) or associative neural networks (implicit), in both of which associative knowledge is represented in a non-action-centered way.

As mentioned before, in a similar fashion but orthogonally, implicitness/explicitness is also distinguished based on representation. Implicit memory can be represented using connectionist distributed representation (such as in the hidden layer of a backpropagation network), which is (relatively speaking) less accessible to cognitive agents possessing it (Sun, 2002, 1999), whereas explicit memory can be represented using symbolic/localist representation, which is relatively more accessible to cognitive agents possessing it (Kirsh, 1990). Implicit and explicit memory thus reside in two different modules with different representations. Moreover, in this way, the two dichotomies are completely separate from each other: That is, there are both implicit and explicit procedural (action-centered) memory, and both implicit and explicit declarative (non-action-centered) memory.

This four-way division is functional in accordance with the ecological-functional perspective, because of (1) the division of labor between explicit and implicit memory (while one is for storing explicit information that is more crisp, the other is for storing implicit information that is more complex), and (2) the division of labor between declarative and procedural memory (while one is for storing general information, the other is for storing information oriented specifically towards action decision making). The divisions of labor led to both the separation and the interaction of these different types of knowledge. The separation ensures that different types of information may be found separately and thus relatively easily, while the interaction among different types of memory helps to bring together different types of information when needed (Klein et al., 2002; Sun & Zhang, 2004, 2007; Sun et al., 2009), to ensure better performance and even "synergy" as mentioned before and as will be discussed later (see Sections 3.2.1 and 3.2.2; see also Sun et al., 2001, 2005). Furthermore, the separation allows different memories to work on different tasks possibly simultaneously and thus enhances the overall functionality. Thus, the four-way division is in keeping with the ecological-functional considerations.

The orthogonality of the distinction between procedural and declarative memory and the distinction between implicit and explicit memory will be further argued for next, along with the issue of semantic versus episodic memory.

3.1.3. *Evidence in support of hypothesis III*

I now turn to the distinction between episodic and semantic memory within the NACS of CLARION, and its orthogonality with the implicit–explicit distinction.

First of all, we look into the distinction between episodic and semantic memory. Quillian (1968) originally proposed the idea of semantic memory for the sake of storing and organizing information for semantic processing (including information about word meanings and information about the world in general). However, this notion has been generalized to include all general knowledge that is not directly related to specific experiences (that is, not episodic in nature) and not action-centered (i.e., not procedural). Tulving (1972, 1983), for example, expounded on the difference between semantic and episodic memory. See Norman, Detre, and Polyn (2008) and Rogers (2008) for various models of semantic and episodic memory respectively.

In line with the ecological-functional perspective (Sun, 2002), Klein et al. (2002) pointed out that "episodic and semantic memory systems evolved to solve many different problems and that they are accessed by many different kinds of decision rules". However, whereas some tasks may require information from episodic memory alone or from semantic memory alone, other tasks may require information from both memory systems (Klein et al., 2002). Dissociations between memory systems may not be absolute; one may find independence for some tasks and joint activation for others. The extent to which one finds functional independence between memory systems may reflect the informational requirements of the activities that the cognitive agent engages in. The division of episodic and semantic memory can thus be functional.

To establish the orthogonality of the implicit-explicit distinction and the semantic-episodic distinction, first, the distinction between implicit and explicit semantic memory needs to be argued. On the one hand, explicit semantic memory is well established. Ever since the days of Quillian (1968), semantic memory has been portrayed as largely explicit and conceptual, consisting of explicit concepts and conceptual relations among them. Collins and Loftus (1975), for instance, advocated such a view. On the other hand, implicit semantic memory may require some explanations. What distinguishes implicit semantic memory from explicit semantic memory is that implicit semantic memory involves implicit connections among memory contents (whereby the existence of these connections is outside of conscious awareness). One notion that is important for implicit semantic memory is priming. As defined by Tulving (1985), priming is a mechanism designed to facilitate the identification of the same or similar object(s) on a subsequent occasion, in the sense that the identification of the object(s) requires less stimulus information or occurs more quickly than it does in the absence of priming (Nelson, McKinney, Gee, & Janczura, 1998). Priming often occurs in the absence of conscious awareness—voluminous data from implicit memory research point to this interpretation (see, e.g., Roediger, 1990; Schacter, 1987; Toth et al., 1994; and so on). Furthermore, Tulving and Schacter (1990) suggested that conceptual priming (including conceptual priming without conscious awareness) involved semantic memory, hence indicating the possibility of implicit semantic memory. It is reasonable to assume that the implicit semantic memory is separate from its explicit counterpart, on the basis of many dissociations between explicit and implicit memory tests (that is, the fact that certain priming affects certain implicit tests but not explicit tests). Such dissociations (especially the "reversed dissociations" discussed in the literature) suggest the possibility of separate implicit and explicit semantic memory systems (Dunn & Kirsner, 1988). See also other arguments earlier regarding the separation of implicit and explicit memory. There have certainly been debates regarding whether dissociations point to different processes or difference systems (see, e.g., Hintzman, 1990; Schacter, Wagner, & Buckner, 1998). However, as pointed out by Hintzman (1990), "once the model has been spelled out, it makes little difference whether its components are called systems, modules, processes, or something else; the explanatory burden is carried by the nature of the proposed mechanisms and their interactions, not by what they are called" (p. 121).

Note that in Schacter's (1987) scheme, there were multiple co-existing implicit memory systems, some of which were declarative and semantic according to the CLARION framework. As added support for this view, Sun and Zhang (2007) showed how the division of implicit and explicit semantic memory might account for categorical inferences where similarity-based processes played a significant role; Hélie and Sun (2010) showed in detail how the division of implicit and explicit semantic memory might explain the processes of creative problem solving.

Second, to establish the orthogonality of the explicit-implicit distinction and the episodic-semantic distinction, the distinction between implicit and explicit episodic memory also needs to be addressed. Explicit episodic memory is, relatively speaking, well established (Tulving, 1983). It stores information concerning prior experiences in an explicit and individuated form. It includes spatial and temporal information about events and activities. It constitutes an explicit personal memory (or "self-referential memory"; Tulving 1983). On the other hand, implicit episodic memory may be used for keeping statistics extracted from actual experiences stored in explicit episodic memory (e.g., Hasher & Zacks, 1979). It is a derived memory system in the sense defined by Klein et al. (2002), which is formed through transforming available information in a way that enables rapid supply of needed information in appropriate circumstances and reduces or even eliminates further processing, so that rapid decision making is made possible. Some may argue that implicit episodic memory should be categorized as a semantic memory, but such a debate would not be very meaningful as it would merely be about a label. For example, what was termed "semantic trait memory" by Klein et al. (2002) is an implicit episodic memory according to the CLARION framework. In general, implicit episodic memory in CLARION can be understood in this light.

In line with the ecological-functional perspective, according to Klein et al. (2002), derived memory is formed on the basis of (a) predictability—it should lead to successful anticipation of likely future use of information, (b) importance—the cost of derived memory should be justified by the value of the prediction, (c) urgency—the information may need to be accessed in a rapid fashion, (d) economy—the number of judgments supported by the derived memory should be large enough compared with the size of the memory (Klein et al., 2002; Schneider, 1993). Pre-stored summaries in implicit episodic memory reduce on-line computation and therefore speed up retrieval, but they require additional representational structures and extra storage. Memory systems may have evolved to address such tradeoffs. The division of implicit and explicit episodic memory is functional in this sense.

3.2. Accounting for memory-related cognitive phenomena

Below I show how the framework of, as well as the perspective embodied by, the CLARION cognitive architecture enables us to account for a number of interesting memory phenomena qualitatively (i.e., conceptually).

3.2.1. Accounting for synergy between implicit and explicit procedural memory

The CLARION framework predicts that there may be a synergy between implicit and explicit procedural (action-centered) memory, resulting from their interaction in learning and in action decision making (see Sun et al., 2005 for details). From an ecological-functional perspective, there is no surprise that the interaction of these two memory systems may be synergistic. We naturally expect that this division of memory (between implicit and explicit procedural memory) should be functional and thus expect some benefits. Judging from the empirical literature, such a synergy may show up, under right circumstances, by speeding up skill learning, improving learned skill performance, and facilitating transfer of learned skills.

There is some empirical evidence in support of this prediction (Sun et al., 2005). In terms of speeding up learning, Willingham et al. (1989) found that those participants in serial reaction time tasks who acquired more explicit knowledge (in explicit procedural memory) appeared to learn faster. It appeared that their explicit procedural memory supplemented their implicit procedural memory. Stanley et al. (1989) reported that, in a process control task, participants' learning improved if they were asked to generate verbal instructions for other participants during learning. That is, a participant was able to speed up his/her own learning through an explication process that generated explicit procedural knowledge (in explicit procedural memory; in addition to implicit procedural knowledge in implicit procedural memory). Sun et al. (2001) showed a similar effect of verbalization in a minefield navigation task, and Reber and Allen (1978) in an artificial grammar learning task. Mathews et al. (1989) showed that a better performance could be attained if a proper mix of implicit and explicit learning was used (in their case, through an experimental condition in which first implicit learning and later explicit learning was encouraged).

In addition, in terms of learned skill performance, Stanley et al. (1989) found that participants who verbalized while performing process control tasks were able to attain a higher level of performance than those who did not verbalize, likely because the requirement that they verbalized their knowledge prompted the formation and utilization of explicit knowledge (from explicit procedural memory), which supplemented their implicit knowledge (from implicit procedural memory). Sun et al. (2001) also showed that verbalizing participants were able to attain a higher level of performance in a minefield navigation task. Squire and Frambach (1990) reported that, initially, amnesic and normal subjects performed comparably in a process control task and equally lacked explicit knowledge. However, with more training, normals achieved better performance than amnesics and also better scores on explicit knowledge measures, which pointed to the possibility that it was because normal subjects were able to learn better explicit knowledge (in explicit procedural memory) that they achieved better performance. Consistent with this interpretation, Estes (1986) suggested that implicit learning alone could not lead to optimal levels of performance. Even in high-level skill acquisition domains, similar effects were observed. Gick and Holyoak (1980) found that good problem solvers could better state explicit rules that described their actions in problem solving. Bower and King (1967) showed that verbalization improved performance in classification rule learning. Gagne and Smith (1962) showed the same effect of verbalization in learning to solve Tower of Hanoi.

In terms of facilitating transfer of learned skills, Willingham et al. (1989) showed some suggestive evidence that explicit procedural knowledge facilitated transfer of learned skills. They reported that (1) participants who acquired explicit knowledge (in explicit procedural memory) in a training task tended to have faster response times in a transfer task; (2) these participants were also more likely to acquire explicit knowledge in the transfer task; and (3) these participants who acquired explicit knowledge responded more slowly when the transfer task was unrelated to the training task (suggesting that the explicit knowledge of the previous task might have interfered with the performance of the transfer task). Sun et al. (2001) showed similar effects. In high-level domains, Ahlum-Heath and DiVesta (1986) found that the participants who were required to verbalize while solving Tower of Hanoi performed better on a transfer task after training than the participants who were not required to verbalize.

Synergy effects are of course dependent on contexts, and not universal (Sun, 2002). Under some circumstances, explicit processes might even hurt performance. Even so, it should be recognized that explicit processes based on explicit procedural memory play important cognitive functions as discussed above. Explicit processes also serve additional functions, such as facilitating verbal communication, or acting as gatekeepers (e.g., enabling conscious veto, as suggested by Libet, 1985).

It has been amply demonstrated (see, e.g., Sun et al., 2005, 2007) that CLARION can account for the synergy effects described above, through the interaction between implicit and explicit procedural memory. The synergy effects may be discerned, in various tasks, through comparing learning conditions, such as comparing the verbalization condition and the non-verbalization condition (whereby the verbalization condition encourages explicit processes), or comparing the dual-task condition and the single-task condition (whereby the dual-task condition discourages explicit processes; see detailed justifications in Sun et al., 2005).

3.2.2. Accounting for synergy between implicit and explicit declarative memory

Similar to the synergy effects between implicit and explicit procedural (action-centered) memory, the CLARION framework predicts that there are also synergy effects between implicit and explicit declarative (non-action-centered) memory. Such synergy effects are expected from an ecological-functional perspective: Like other divisions of memory, this division should be functional also.

Let us examine some experiments with artificial grammar learning (Reber, 1989). Recall that in process control experiments, synergistic results were found between implicit and explicit procedural memory during task performance (see the previous subsection and also Sun et al., 2005). Domangue, Mathews, Sun, Roussel, and Guidry (2004) investigated the effects of similar training variables in artificial grammar learning, examining these effects in situations involving implicit and explicit declarative (as opposed to procedural) memory. The series of experiments in Domangue et al. (2004) (also Sun & Mathews, 2005) examined implicit training (encouraging implicit processes), explicit training (encouraging explicit processes), mixing training across sessions, as well as integrated training providing simultaneous experience with exemplars (encouraging implicit processes) and the grammar (encouraging explicit processes).

The results followed the pattern that encouraging explicit processing generally led to slower but more accurate responses on the cued-generate test. Encouraging implicit processing led to faster responses but with lower accuracy. In contrast, the integrated training achieved a proper balance: having higher accuracy than implicit training, and higher speed than explicit training. One reasonable interpretation within the CLARION framework is that explicit training in this particular context led to the encoding of more grammatical knowledge in the form of explicit rules in explicit declarative memory, while implicit training led to the encoding of more implicit associative mappings in implicit declarative memory. This was because neither implicit nor explicit knowledge in this task was action-centered—they likely resided in declarative memory (explicit or implicit). Often, both kinds of learning occurred, and the differences among different training conditions lay in the proportions of, and the interactions between, the two kinds.

The experimental results above were simulated using CLARION, based on the interpretation stated above. The simulation correspondingly demonstrated the synergy effect of the integrated training, which achieved higher accuracy than implicit training and higher speed than explicit training. See Sun and Mathews (2005) for details.

There is also other corroborating evidence pointing to the synergy between implicit and explicit declarative memory. For example, Berry (1983) showed in a reasoning task that verbalization during learning improved transfer performance. The result appeared to indicate the synergy between implicit and explicit declarative memory (as well as between implicit and explicit reasoning on their basis). Nokes and Ohlsson (2001) showed related results as well. This phenomenon may also be related, to some extent, to the self-explanation effect reported in the literature (e.g., Chi, Bassok, Lewis, Reimann, & Glaser, 1989): Participants who explained examples in physics textbooks more completely did better in solving new problems. There are also indications of positive effects of interaction (i.e., synergy effects) in alphabetic arithmetic tasks, categorical inference tasks, and so on. In all these cases, it could be the use of explicit declarative knowledge (explicit declarative memory), in addition to the use of implicit declarative knowledge (implicit declarative memory), that helped the performance. CLARION has been used to capture data and phenomena in some of these tasks as well (see, e.g., Sun & Zhang, 2007; Sun, Zhang, & Mathews, 2009).

In addition, similarity-based reasoning can be naturally carried out through the interaction of implicit and explicit declarative memory. The interaction of the two kinds of memory within the NACS of CLARION is so structured that the activation flows (from one to the other and then back) account for important kinds of similarity-based processes in human everyday reasoning (see Sun, 1994, 2003 for details). Furthermore, the combination of such similarity-based reasoning and rule-based reasoning (carried out with implicit and explicit declarative memory) may capture and explain a wide range of common human everyday reasoning patterns (as shown by Sun, 1994 and Sun & Zhang, 2007). Thus, the separation and the interaction of these two kinds of memory are highly functional.

In sum, there is evidence that the division between implicit and explicit declarative memory is functional in an ecological-functional sense (e.g., the synergy effects), and CLARION was able to capture and explain the effects of this division.

3.3. A critique of alternative views

There have been other integrated cognitive models (including cognitive architectures) that contain multiple memory systems, and they address the relations and interactions among memory systems (modules). For example, Anderson's cognitive architectures, ACT* and ACT-R (Anderson, 1983; Anderson & Lebiere, 1998), belong to this category.

ACT* consists of a production (rule-based) system and a semantic network. The semantic network represents declarative memory, and the production system represents procedural memory. The distinction between short-term and long-term memory is reduced to activation traces as short-term memory and established nodes and links as long-term memory. The distinction of semantic and episodic memory is not dealt with. As mentioned before, in Anderson (1983), declarative knowledge is assumed to be explicit: Subjects can manipulate and report on such knowledge. Procedural knowledge is not: It leads to actions without much accessibility. In ACT-R (as described by Anderson & Lebiere, 1998), each individual piece of knowledge involves both subsymbolic and symbolic representations. One interpretation is that the symbolic representation is explicit while the subsymbolic representation is implicit (for declarative knowledge, or for both declarative and procedural knowledge).

As discussed before, the first view above unnecessarily confounds two aspects: action-centeredness and accessibility. This is because action-centeredness does not necessarily go with implicitness, and non-action-centeredness does not necessarily go with explicitness (as discussed earlier). The alternative view that each individual piece of knowledge involves both implicit and explicit representations is also problematic, because this claim contradicts the phenomenon that some knowledge may be completely implicit (as discussed earlier). In contrast, CLARION is based on the separation of the two dichotomies, as there are both implicit and explicit procedural memory, and both implicit and explicit declarative memory (see the detailed arguments earlier). In addition, in CLARION, while activations of rules and chunks may be viewed as short-term memory, there is also a dedicated working memory for the sake of facilitating action decision making (see explanations and justifications before; see Sun, 2003 for details).

There have also been some models combining implicit and explicit processes. For example, Cleeremans (1994) used a simple buffer network to capture the effect of explicit knowledge, along with a simple recurrent network for capturing implicit knowledge. However, in more complex tasks, the buffer network, as is, may be inadequate for capturing explicit knowledge used in performing these tasks (e.g., see Sun et al., 2001). Furthermore, in more complex tasks, implicit and explicit knowledge may have more complex interactions. While CLARION can accommodate more complex interactions, Cleeremans' model may not. The same may be said about Erickson and Kruschke's (1998) model. Notably, beside CLARION, there is thus far no other model that is full compatible with the review of empirical evidence given earlier.

4. General discussion

In the preceding sections, I have examined a large body of empirical literature, and compared the memory systems of CLARION to that literature. The general conclusion was that there was substantial empirical support for the CLARION framework. However, it is rarely the case that empirical evidence points unambiguously to a theory (or model). It depends on many other factors (for detailed discussions, see Sun, 2009). The ability to model empirical phenomena (qualitatively or quantitatively) by a theory/model is a necessary, but not sufficient, condition for validity.

In addition to what has been discussed so far, various simulations have been carried out that provide support for the CLARION framework. For example, for validating the distinction between implicit and explicit procedural memory, serial reaction time tasks, process control tasks, alphabetic arithmetic tasks, a complex minefield navigation task, Tower of Hanoi, and other tasks have been simulated within CLARION. For validating the distinction between implicit and explicit declarative memory, categorical inference tasks, incubation tasks, insight problem solving tasks, artificial grammar learning tasks, similarity-based reasoning, and so on have been simulated. Some of the tasks also pointed to the orthogonality of the two dichotomies (implicit versus explicit and procedural versus declarative). These studies have been reported in various previous publications (e.g., Hélie & Sun, 2010; Sun et al., 2001, 2005; Sun & Zhang, 2004, 2007; Sun, Zhang, Slusarz, & Mathews, 2007; Sun et al., 2009), and therefore will not be repeated here. Note that, instead of choosing typical "memory tasks", we chose to use more general psychological tasks in our simulation of the effects of different memory systems and their interactions. The reason was that we chose to focus on how memory fits into the everyday activities of cognitive agents and the role of memory in the overall functioning of cognitive agents. This is our non-memory-centered way of looking at memory.

However, is it possible that alternative frameworks of memory systems can be developed and argued for based on empirical and theoretical considerations comparable to those summarized in this article? It is certainly possible. There is no a priori reason to believe that alternative frameworks can be easily dismissed. However, if one looks at the totality of my arguments on the basis of both theoretical considerations (including those in Sun, 2002, 2004) and empirical evidence, my proposed framework should be a strong contender among other possible proposals regarding the overall organization of memory systems. There is currently no other framework that is as comprehensive and well-supported empirically by a wide range of data and simulations (and I am ready to consider any emerging alternatives when they are sufficiently developed).

In this article, I focus on the overall architecture of memory systems, not individual aspects of memory or specialized models of memory. There have been many specialized memory models proposed (e.g., Erickson & Kruschke, 1998; Hintzman, 1986; Metcalfe, 1991; see Norman et al., 2008 and Rogers, 2008 for reviews). It is also worth noting that our work is a broad-stroke interpretation of empirical work, and the CLARION framework is not a strict derivation from existing empirical data. My approach starts out with an analysis of memory functions within the context of everyday activities (Sun, 2002, 2004). The result is a specification of the organization of memory. The framework succeeded in accounting for a range of empirical phenomena and data. However, it is important to note that data matching is not a starting point, but a result of the analysis.

In all, a major motivation driving this work is that there is the need to address the issue of an overall framework of memory systems from the standpoint of cognitive architectures. In this article, a synthesis of memory systems within a cognitive architecture is presented. Although the cognitive architecture has been described before (in bits and pieces in psychological and cognitive science journals), an overview and a synthesis as in the present paper are very much needed. My framework included implicit procedural memory, explicit procedural memory, implicit declarative memory (including implicit semantic and episodic memory), explicit declarative memory (including explicit semantic and episodic memory), as well as auxiliary memory modules. It constituted a set of hypotheses from the perspective of a cognitive architecture. CLARION shows clearly how they fit together in a mechanistic, process-based way: CLARION reconciles various distinctions to come up with an overall architecture of the mind. In this article, I explored, in detail, how these hypotheses might be validated through examining the empirical literature on memory and beyond. Furthermore, some preliminary evaluation of the hypotheses was conducted through computational simulations of a variety of quantitative data based on the CLARION cognitive architecture. Results of both qualitative and quantitative data matching point to the promise of the proposed framework. Generally I stressed the need for an examination of memory systems from a broad, non-memory-centered perspective.

Acknowledgment

This work has been supported in part by Army Research Institute grant W74V8H-05-K-0002 (to Ron Sun and Bob Mathews) and Office of Naval Research grant N00014-08-1-0068 (to Ron Sun). Thanks to Robert Mathews, Xi Zhang, Sebastien Helie, and others for discussions.

References

Ackerman, P., & Kanfer, R. (2004). Cognitive, affective, and conative aspects of adult intellect within a typical and maximal performance framework. In D. Y. Dai, & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition* (pp. 119–141). Mahwah, NJ: Lawrence Erlbaum Associates.

Ahlum-Heath, M., & DiVesta, F. (1986). The effect of conscious controlled verbalization of a cognitive strategy on transfer in problem solving. *Memory and Cognition, 14*, 281–285.

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

Anderson, J., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates.

Baddeley, A. (1986). *Working memory*. New York: Oxford University Press.

Berry, D. (1983). Metacognitive experience and transfer of logical reasoning. *Quarterly Journal of Experimental Psychology, 35A*, 39–49.

Berry, D., & Broadbent, D. (1988). Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology, 79*, 251–272.

Bower, A., & King, W. (1967). The effect of number of irrelevant stimulus dimensions, verbalization, and sex on learning biconditional classification rules. *Psychonomic Science, 8*(10), 453–454.

Bruner, J., Goodnow, J., & Austin, J. (1956). *A study of thinking*. New York: Wiley.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual process theories in social psychology*. New York: Guilford Press.

Clark, A. (1997). *Being there*. Cambridge, MA: MIT Press.

Cleeremans A., (1994). Attention and awareness in sequence learning. Proc. of Cognitive Science Society Annual Conference, 330–335.

Cleeremans, A. (1997). Principles for implicit learning. In D. Berry (Ed.), *How implicit is implicit learning?* (pp. 195–234). Oxford, UK: Oxford University Press.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: news from the front. *Trends in Cognitive Sciences, 2*(10), 406–416.

Collins, A., & Loftus, J. (1975). Spreading activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Coward, L. A., & Sun, R. (2004). Criteria for an effective theory of consciousness and some preliminary attempts. *Consciousness and Cognition, 13*, 268–301.

Chi, M., Bassok, M., Lewis, M., Reimann, P., & Glaser, P. (1989). Self-explanation: how students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145–182.

Domangue, T., Mathews, R., Sun, R., Roussel, L., & Guidry, C. (2004). The effects of model-based and memory-based processing on speed and accuracy of grammar string generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(5), 1002–1011.

Dreyfus, H. (1992). *Being-in-the-world*. Cambridge, MA: MIT Press.

Dreyfus, H., & Dreyfus, S. (1987). *Mind over machine: The power of human intuition*. New York: The Free Press.

Dunn, J., & Kirsner, K. (1988). Discovering functionally independent mental processes: the principle of reversed associations. *Psychological Review, 95*, 21–101.

Erickson, M., & Kruschke, J. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127*, 107–140.

Estes, W. (1986). Memory storage and retrieval processes in category learning. *Journal of Experimental Psychology: General, 115*, 155–174.

Evans, J., & Frankish, K. (Eds.). (2009). *Two minds: Dual processes and beyond*. Oxford, UK: Oxford University Press.

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Gagne, R., & Smith, E. (1962). A study of the effects of verbalization on problem solving. *Journal of Experimental Psychology, 63*, 12–18.

Gick, M., & Holyoak, K. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306–355.

Hasher, J., & Zacks, J. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*, 356–358.

Heidegger, M. (1927). *Being and time*. English translation published by Harper and Row, New York. 1962.

Hélie, S., & Sun, R. (2010). Incubation, insight, and creative problem solving: a unified theory and a connectionist model. *Psychological Review, 117*, 994–1024.

Hintzman, D. (1986). Schema abstraction in a multiple-trace memory model. *Psychological Review, 93*, 411–428.

Hintzman, D. (1990). Human learning and memory. *Annual Review of Psychology, 41*, 109–139.

Hirschfield, L., & Gelman, S. (Eds.). (1994). *Mapping the mind: Domain specificity in cognition and culture*. Cambridge, UK: Cambridge University Press.

Jacoby, L. (1983). Remembering the data: analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior, 22*, 485–508.

Kanfer, R., & Ackerman, P. (1989). Motivation and cognitive abilities: an integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*(4), 657–690.

Kirsh, D. (1990). When is information explicitly represented. In P. Hanson (Ed.), *Information, language, and cognition*. Vancouver, Canada: University of British Columbia Press.

Klein, S., Cosmides, L., Tooby, J., & Chance, S. (2002). Decisions and the evolution of memory: multiple systems, multiple functions. *Psychological Review, 109*(2), 306–329.

LeDoux, J. (1996). *The emotional brain*. New York: Simon and Schuster.

Lewicki, P., Czyzewska, M., & Hoffman, H. (1987). Unconscious acquisition of complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13*(4), 523–530.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences, 8*, 529–566.

Logan, G. (2002). An instance theory of attention and memory. *Psychological Review, 109*(2), 376–400.

McClelland, J., McNaughton, B., & O'Reilly, R. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419–457.

Mathews, R., Buss, R., Stanley, W., Blanchard-Fields, F., Cho, J., & Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: a synergistic effect. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 1083–1100.

Metcalfe, J. (1991). Recognition failure and the composite memory trace in CHARM. *Psychological Review, 98*(4), 529–553.

Moscovitch, M., & Umilta, C. (1991). Conscious and unconscious aspects of memory. In *Perspectives on cognitive neuroscience*. New York: Oxford University Press.

Nelson, D., McKinney, V., Gee, N., & Janczura, G. (1998). Interpreting the influence of implicitly activated memories on recall and recognition. *Psychological Review, 105*(2), 299–324.

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

Nisbett, R., & Wilson, T. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review, 84*(3).

Norman, K., Detre, G., & Polyn, S. (2008). Computational models of episodic memory. In R. Sun (Ed.), *Cambridge handbook on computational psychology*. New York: Cambridge University Press.

Nokes, T. J., & Ohlsson, S. (2001). How is abstract generative knowledge acquired? A comparison of three learning scenarios. In J. D. Moore, & K. Stenning (Eds.), *Proceedings of the twenty third annual conference of the cognitive science society* (pp. 710–715). Mahaw, N.J.: Erlbaum.

Proctor, R., & Dutta, A. (1995). *Skill acquisition and human performance*. Thousand Oaks, CA.: Sage Publications.

Quillian, M. R. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic information processing* (pp. 227–270). Cambridge, MA: MIT Press.

Ratcliff, R., & McKoon, G. (1998). Memory models. In E. Tulving (Ed.), *Oxford handbook of memory* (pp. 571–581). New York: Oxford University Press.

Reber, A. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*(3), 219–235.

Reber, A., & Allen, R. (1978). Analogy and abstraction strategies in synthetic grammar learning: a functionalist interpretation. *Cognition, 6*, 189–221.

Roediger, H. (1990). Implicit memory: retention without remembering. *American Psychologist, 45*(9), 1043–1056.

Rogers, T. (2008). Computational models of semantic memory. In R. Sun (Ed.), *Cambridge handbook on computational psychology*. New York: Cambridge University Press.

Rosenbaum, D., Carlson, R., & Gilmore, R. (2001). Acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology, 52*, 453–470.

Rosenbloom, P., Laird, J., & Newell, A. (1993). *The SOAR papers: Research on integrated intelligence*. Cambridge, MA: MIT Press.

Rumelhart, D., McClelland, J., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructures of cognition*. Cambridge, MA: MIT Press.

Schacter, D. (1987). Implicit memory: history and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 501–518.

Schacter, D., Wagner, A., & Buckner, R. (1998). Memory systems of 1999. In E. Tulving (Ed.), *Handbook of memory* (pp. 627–643). New York: Oxford University Press.

Schneider, W. (1993). Varieties of working memory as seen in biology and in connectionist/control architectures. *Memory and Cognition, 21*(2), 184–192.

Seger, C. (1994). Implicit learning. *Psychological Bulletin, 115*(2), 163–196.

Shanks, D., & St.John, M. (1994). Characteristics of dissociable learning systems. *Behavioral and Brain Sciences, 17*, 367–394.

Squire, L. (1987). *Memory and brain*. New York: Oxford University Press.

Squire, L., & Frambach, M. (1990). Cognitive skill learning in amnesia. *Psychobiology, 18*, 109–117.

Stanley, W., Mathews, R., Buss, R., & Kotler-Cope, S. (1989). Insight without awareness: on the interaction of verbalization, instruction and practice in a simulated process control task. *Quarterly Journal of Experimental Psychology, 41A*(3), 553–577.

Sun, R. (1994). *Integrating rules and connectionism for robust commonsense reasoning*. New York: John Wiley and Sons.

Sun, R. (1999). Accounting for the computational basis of consciousness: a connectionist approach. *Consciousness and Cognition, 8*, 529–565.

Sun, R. (2002). *Duality of the mind*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sun, R. (2003). *A tutorial on CLARION*. Technical report. Cognitive Science Department, Rensselaer Polytechnic Institute. http://www.cogsci.rpi.edu/~rsun/sun.tutorial.pdf.

Sun, R. (2004). Desiderata for cognitive architectures. *Philosophical Psychology, 17*(3), 341–373.

Sun, R. (2009). Motivational representations within a computational cognitive architecture. *Cognitive Computation, 1*(1), 91–103.

Sun, R., & Mathews, R. (2005). *Exploring the interaction of implicit and explicit processes to facilitate individual skill learning*. Technical Report TR-1162. Arlington, VA: Army Research Institute for the Social and Behavioral Sciences.

Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge: a bottom-up model of skill learning. *Cognitive Science, 25*(2), 203–244.

Sun, R., Slusarz, P., & Terry, C. (2005). The interaction of the explicit and the implicit in skill learning: a dual-process approach. *Psychological Review, 112*(1), 159–192.

Sun, R., & Zhang, X. (2004). Top-down versus bottom-up learning in cognitive skill acquisition. *Cognitive Systems Research, 5*(1), 63–89.

Sun, R., & Zhang, X. (2007). Accounting for a variety of reasoning data within a cognitive architecture. *Journal of Experimental and Theoretical Artificial Intelligence, 19*(2), 159–193.

Sun, R., Zhang, X., & Mathews, R. (2009). Capturing human data in a letter counting task: accessibility and action-centeredness in representing cognitive skills. *Neural Networks, 22*, 15–29.

Sun, R., Zhang, X., Slusarz, P., & Mathews, R. (2007). The interaction of implicit learning, explicit hypothesis testing learning, and implicit-to-explicit knowledge extraction. *Neural Networks, 20*(1), 34–47.

Toth, J., Reingold, E., & Jacoby, L. (1994). Toward a redefinition of implicit memory: process dissociations following elaborative processing and self-generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(2), 290–303.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving, & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: AcademicPress.

Tulving, E. (1983). *Elements of episodic memory*. Oxford: Clarendon Press.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40*, 385–398.

Tulving, E., & Schacter, D. (1990). Priming and human memory systems. *Science, 247*, 301–305.

Vicente, K., & Wang, J. (1998). An ecological theory of expertise effects in memory recall. *Psychological Review, 105*(1), 33–57.

Warrington, E., & Weiskrantz, L. (1970). Amnesic syndrome: consolidation or retrieval. *Nature, 217*, 972–974.

Watkins, C. (1989). Learning with Delayed Rewards. Ph.D Thesis, Cambridge University, Cambridge, UK.

Willingham, D. (1998). A neuropsychological theory of motor skill learning. *Psychological Review, 105*(3), 558–584.

Willingham, D., Nissen, M., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1047–1060.