Psychological Review
2002, Vol. 109, No. 1, 55–65

0033-295X/02/$5.00 DOI: 10.1037//0033-295X.109.1.55

A Hybrid Architecture for Working Memory: Reply to MacDonald and Christiansen (2002)

Marcel Adam Just and Sashank Varma
Carnegie Mellon University

This article responds to M. C. MacDonald and M. H. Christiansen's 2002 commentary on the capacity theory of working memory (WM) and its computational implementation, the Capacity-Constrained Collaborative Activation–based Production System (3CAPS). The authors also point out several shortcomings in MacDonald and Christiansen's proposal for the construal of WM, arguing that at some level of description, their model is a variant of a small subset of the 3CAPS theory. The authors go on to describe how the symbolic and connectionist mechanisms within the hybrid 3CAPS architecture combine to produce a processing style that provides a good match to human sentence comprehension and other types of high-level cognition. The properties of 3CAPS are related to the development of other connectionist, symbolic, and hybrid systems.

This article has the goals of (a) refuting some of MacDonald and Christiansen's (2002) incorrect descriptions of the capacity theory of sentence comprehension as described in Just and Carpenter (1992); (b) pointing out the theoretical and empirical difficulties with MacDonald and Christiansen's alternative approach and with their simple recurrent network (SRN) model in particular; and (c) pointing out some commonalities between symbolic, connectionist, and hybrid approaches, describing some of their formal properties and their application to a number of cognitive phenomena.

Misconstruals in MacDonald and Christiansen's (2002) Rendition of the Capacity Theory

There are several issues on which MacDonald and Christiansen have attributed to Just and Carpenter (1992) a position that was never taken. MacDonald and Christiansen constructed an implausible straw man and then proceeded to argue against it and claim for themselves a sensible high ground. For example, they claimed for themselves the position that comprehension performance is conjointly determined by biological and experiential factors and attributed to Just and Carpenter some unspecified alternative position, such as treating working memory (WM) capacity as a "primitive." We will point out that on many issues, the original Just and Carpenter position or its successors provide a far richer theoretical framework than anything MacDonald and Christiansen offer.

The Putative Separation Between WM and Knowledge

MacDonald and Christiansen inaccurately described the Just and Carpenter (1992) position, attributing to it a separation between WM and the procedural knowledge used in language processing. Contrary to MacDonald and Christiansen's attribution, the capacity theory claims that the procedural knowledge underlying comprehension, along with the activation resources, is part of the WM system, as the following quotations from Just and Carpenter's 1992 article indicate:

> In our theory, working memory for language refers to a set of processes and resources that perform language comprehension. (p. 123)
>
> Of particular relevance are the processes that perform language comprehension. These processes, in combination with the storage resources, constitute working memory for language. (p. 123)

The processes referred to are the productions that embody the procedural knowledge that is used in language comprehension. During the operation of the CC READER model, the functions of the resources and the productions are completely intertwined. In this dynamic system, the productions cannot do anything without activation, and the activation resource is meaningless without the productions. Knowledge and resources are not separable, in this critical sense.

When MacDonald and Christiansen evaluated whether knowledge and resources are separable in connectionist models, they used precisely this sense of separability, namely that ". . . these manipulations [of capacity in connectionist models, such as the manipulation of the number of hidden units, amount of training, efficiency, or amount of noise] affect the behavior of the whole network, both its processing and its representation" (p. 38). Capacity-related attributes can vary independently of linguistic knowledge in the connectionist models that MacDonald and Christiansen cited. For example, increasing the noise in the input signal manipulates the representational quality independent of the grammatical knowledge of the network (i.e., the matrix of connection weights), which remains constant. It is conceptually similar to

---

Marcel Adam Just and Sashank Varma, Department of Psychology, Carnegie Mellon University.

This research was partially supported by Office of Naval Research Grant N00014-96-1-0322, National Institute of Mental Health Grant MH-29617, and National Institute of Mental Health Research Scientist Award MH-00662. We thank Akira Miyake for helpful comments on this article.

Correspondence concerning this article should be addressed to Marcel Adam Just, Center for Cognitive Brain Imaging, Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213. E-mail: just+@cmu.edu

manipulating the capacity of a Capacity-Constrained Collaborative Activation–based Production System (3CAPS) model while holding its knowledge (productions) constant. In both cases, the behavior of the whole system is affected. Thus, substantial portions of the MacDonald and Christiansen article argued about a nonexistent difference.

The Role of Experience in the Capacity Theory

Despite MacDonald and Christiansen's implications to the contrary, the Just and Carpenter (1992) position fully acknowledges the role of experience as a partial determinant of WM capacity, as these quotations indicate:

> As we discussed earlier, the individual differences reported here may reflect differences in total capacity, differences in processing efficiency, or both. . . . In contrast, a change in processing efficiency is assumed to be more specific to a particular process. Thus, changes in the efficiency of a process are often assumed to result from practice or some instructional intervention. Indeed, intensive practice in several simple tasks, such as Stroop-type tasks, induces large changes in the speed of responding that are typically interpreted in terms of changes in efficiency of underlying processes (Cohen, Dunbar & McClelland, 1990; Schneider & Shiffrin, 1977). *Intensive practice in reading might similarly induce greater efficiency in some component processes of comprehension; the time spent in out-of-school reading is correlated with reading skill in fifth-grade students, accounting for approximately 9% of the variance in one study* [italics added] (Anderson, Wilson, & Fielding, 1988). (Just & Carpenter, 1992, p. 145)
>
> However, another account of individual differences is in terms of the efficiency of mental processes. . . . The two explanations are mutually compatible and the experiments described here do not attempt to discriminate between them. (Just & Carpenter, 1992, p. 124)

Thus, MacDonald and Christiansen incorrectly characterized the Just and Carpenter position, which explicitly recognized the relationship between experience and WM capacity.

Caplan and Waters (2002)

As part of their reply to MacDonald and Christiansen, Caplan and Waters questioned some of the Just and Carpenter (1992) and King and Just (1991) results regarding individual differences in various sentence-comprehension tasks. Caplan and Waters's central assertion (they claimed that there is no exacerbation of the processing difference between high- and low-span readers at points of peak syntactic processing load) can be evaluated against the several sources of data summarized by Just, Carpenter and Keller (1996), showing their assertion to be wrong. Because this questioning has been previously published (Waters & Caplan, 1996) and refuted (Just, Carpenter, & Keller, 1996), we choose not to reproduce the refutation here on the grounds that, in our view, the Caplan and Waters questioning is not scientifically productive, as it was not in 1996.

Comments on the MacDonald and Christiansen (2002) Proposal

MacDonald and Christiansen Failed to Account for Constraint

A key part of the theory that MacDonald and Christiansen failed to capture in their approach is a systematic treatment of capacity *constraint*. They offered no disciplined mechanism to account for the fact that there is simply a human limit on thinking. No matter how much experience one has in a task, there is an upper bound on immediate processing capability. MacDonald and Christiansen's approach did not deal with this real and shaping force in human cognition. The capacity theory provides a formal, systematic account of such constraints and their influence on the shape of processing. For example, the framework of the capacity theory has proved useful in designing and interpreting functional magnetic resonance imaging (fMRI) study (Callicott et al., 1999), where it has been possible to study how the constraint applies in different cortical areas. For example, in an *N*-back task, the constraint on *N* (the total number of sequenced items to be held, compared, and updated) that is observed in behavioral performance (i.e., the largest *N* for which accuracy remains high) corresponds to the constraint on activation in the prefrontal cortex. By contrast, MacDonald and Christiansen's dismissal of a constrained WM provided no insight into these fascinating issues concerning the locus of constraint.

MacDonald and Christiansen Failed to Account for Psychometric Evidence

The tradition of psychometrics constructs tests that are intended to measure a particular ability and observes the correlations between performance on those tests and on other tasks that are presumed to draw on that same ability. In this spirit, the reading-span test was constructed and succeeded in correlating with many cognitive tasks, not only straightforward sentence-comprehension tasks. No one has accorded any privileged status to the reading-span task other than its usefulness in predicting comprehension and other performance based on its intended ability to measure WM capacity for language. *Privilege* is not a relevant attribute, but ability to predict performance is. So although the reading-span test is no more privileged than a lexical decision task, as MacDonald and Christiansen (pp. 38–39) suggested, it is immensely more useful in predicting individual differences in comprehension performance. The relevance of this point is that the ability of a psychological theory to account for individual differences is considered a strength of the theory, so the predictive ability of the reading-span test is an additional element of strength of the capacity theory.

Much of MacDonald and Christiansen's facile dismissal of the correlation between the reading-span task and comprehension performance is largely based on their demonstrably false claim that the correlation arises only because the reading-span task is another form of a sentence-comprehension task. The psychometric finding that MacDonald and Christiansen ignored is that there is a substantial correlation between nonlinguistic but symbolic WM tasks (such as math-processing plus storage tasks) and sentence-processing ability, as indicated in Daneman and Merikle's (1996) meta-analysis. This correlation is extremely difficult to explain on the basis of MacDonald and Christiansen's account of differential experience or ability in language, but it is entirely consistent with the capacity theory's conception of an operational capacity.

One example of this type of result that MacDonald and Christiansen could not explain is Roberts and Gibson's (in press) finding of a substantial correlation between nonlinguistic WM measures (such as *N*-back) and sentence-comprehension performance. Roberts and Gibson made their interpretation clearly:

> MacDonald & Christiansen's account, however, addresses only the relation between language processing and linguistic measures of working memory, such as Daneman and Carpenter's (1980) Reading Span. . . . A skill-via-experience account, in which better readers do better in reading comprehension and in linguistic working memory tasks, offers no explanation for correlations among linguistic and non-linguistic working memory tasks, and no explanation for correlations between working memory, as measured by these tasks, and sentence memory.

Roberts and Gibson went on to say that

> this finding [the obtained correlation between non-linguistic measures of WM capacity and sentence memory] casts doubt on MacDonald & Christiansen's (1998) hypothesis that correlations between linguistic working memory measures and sentence comprehension measures are due to the fact that both are sensitive to participants' reading ability. (pp. 13–14)

MacDonald and Christiansen's failure to account for such results has broader implications. As Roberts and Gibson (in press) pointed out,

> These results are also relevant to MacDonald & Christiansen's (1998) attempt to abolish the working memory construct. Their alternative skill-via-experience account, such that better readers are better at both linguistic working memory and sentence comprehension tasks, does not account for the correlations observed in this study. (pp. 23–24)

Roberts and Gibson also say that "memory for sentences is not simply a result of linguistic experience; rather it is likely that there is an independent working memory component contributing to participants' performance on the sentence memory task" (p. 2).

One might also ask why MacDonald and Christiansen's phonological account of individual differences in comprehension ability has not led to a more predictive measure of comprehension performance, perhaps based on differential precision of phonological representation. One of the strengths of the capacity theory is its ability to indicate how individual differences might be measured. It is ironic that MacDonald and Christiansen's account of the difference between high- and low-span readers rests on results obtained by using the very test they question.

MacDonald and Christiansen's hypothesis of the phonological basis of WM differences is an old one to which little new has been added. The hypothesis that individual differences in phonological processing are among the sources of individual differences in reading comprehension was previously proposed by many researchers—for example, the verbal-fluency theory, proposed over two decades ago (Perfetti & Lesgold, 1977), that differences in the efficiency of phonological processing (the "biological" component of MacDonald and Christiansen's account) propagate through all levels of language processing and thus underlie individual differences in comprehension performance. The hypothesis's age did not make it any less true when MacDonald and Christiansen reproposed it, but its age does demonstrate that it does not emanate exclusively from their perspective and in fact predates modern connectionism. However, despite their inflated claim to having encompassed the interaction of biological and experiential factors, this is the main "biological" factor that MacDonald and Christiansen discussed concerning the development of comprehension skill.

It is interesting that in the course of describing the effects of phonological processing in comprehension, MacDonald and Christiansen inconsistently evoked either a capacity explanation or a precision-of-representation explanation. In some cases, they suggested that it is the amount of activation and processing that is the determinant of performance, saying, for example, that "larger extrinsic loads create more phonological activation to compete with the activation needed for sentence processing" (p. 45). This type of account fits a capacity notion, suggesting that there are individual differences in the amount of phonological processing and storage that can be done concurrently. So rather than doing away with WM capacity, MacDonald and Christiansen simply relabeled one facet of WM as *phonological processing* and localized some of the individual differences there. Although their formulation can be construed within a connectionist framework, it can obviously be expressed just as easily without the help of that framework.

Finally, there are populations in which the phonological, word-reading approach makes the wrong predictions. For example, high-functioning autistic readers have a slight advantage over control subjects in reading individual words (Goldstein, Minshew, & Siegel, 1994). From this result, a phonological-processing account would predict that high-functioning autistic subjects should also have an advantage in sentence comprehension. Contrary to this prediction, high-functioning autistic subjects have a disadvantage in comprehending complex verbal instructions. In sum, there is much more to individual differences in sentence comprehension than phonological processing and word reading.

MacDonald and Christiansen Overstated the Adequacy of the Experiential Approach

Although it is plausible that differential frequencies of various language structures contribute to their relative processing difficulty, other important determinants are in play, some of which can dominate frequency effects. Put simply, human cognition is much more than a mirror of the encounters with statistical regularities in the environment. For example, Gibson and Schutze (1999), in describing their evaluation of a frequency-based approach, entitled their article "Disambiguation Preferences in Noun Phrase Conjunction Do Not Mirror Corpus Frequency" and stated that the obtained pattern of results is not "predicted by other exposure-based accounts of ambiguity resolution in sentence comprehension, including lexically based constraint–satisfaction proposals (e.g., MacDonald et al., 1994 . . .) or connectionist-network accounts (Christiansen, 1996 . . .)" (p. 275).

MacDonald and Christiansen's Proposal Is Not Empirically Distinguishable

Before arguing for indistinguishability, we note that the point-to-point correspondence between MacDonald and Christiansen's proposal and 3CAPS is difficult to establish for several reasons. For one thing, the SRN model's account of grammatical category prediction constitutes only a small fraction of the levels of processing and tasks that 3CAPS models account for, so the correspondence is at best a part-to-whole mapping. Another point that is difficult to distinguish empirically is the explanation for the individual differences, which MacDonald and Christiansen postulated, stem in large part from differential amounts of language experience but which Just and Carpenter (1992) took as a point of departure, focusing on characterizing the differences in functioning among individuals of different ability levels. Are the accounts different from each other? MacDonald and Christiansen claimed that they are, and although in some fine-grained ways they do differ, as noted below in the differences in the account of the reading times in subject- and object-relative sentences—alternative formalisms typically have different perspectives or foci. However, in the large, the MacDonald and Christiansen proposal is not different in any important way.

Roberts and Gibson (in press) commented on the indistinguishability of the MacDonald and Christiansen proposal:

> This view [MacDonald & Christiansen, 1998] does not provide a convincing alternative that explains the existing data. First, it is not clear that it is actually an alternative: The explanation for individual differences (differences in language processing skill) translates easily into working memory models such as those of Just and Carpenter (1992), or Salthouse (1990), which view working memory capacity as the interaction of storage capacity and processing efficiency. (p. 13)
>
> They [MacDonald & Christiansen, 1998] state that “maintaining a set of unrelated words requires substantial activation of phonological representations” (p. 14). Maintaining phonological activation of words, however, is another way of describing the storage functions of working memory. Thus, MacDonald and Christiansen have not, in fact, presented an adequate alternative to the idea of variance in working memory capacity as the source of individual differences in understanding language, but have restated the problem. (pp. 23–24)

MacDonald and Christiansen’s Proposal Failed to Take Into Account Brain Imaging Studies of WM

Even though MacDonald and Christiansen claimed to deal with biological factors in language, they remained very distant from any biological mechanism or explanation. Despite unprecedented advances in the past decade in discovering the neural bases of sentence comprehension, MacDonald and Christiansen and their approach have had virtually nothing to say about the biology of sentence comprehension. What are the predictions of their approach concerning the brain activation of better versus poorer comprehenders? Should high-span subjects display more or less activation than low-span subjects, and in which areas; or should individual differences in comprehension be manifested in some other way? What should the difference be in brain activation between the comprehension of more complex and less complex sentences? MacDonald and Christiansen and their approach are mute.

The capacity theory, by contrast, has been among the leading conceptual frameworks for understanding such issues. For example, the capacity theory predicted that more brain activation would be associated with object-relative sentences than subject-relative sentences. The prediction stems from the greater capacity utilization exhibited by the 3CAPS model. The fMRI results (Just, Carpenter, Keller, Eddy, & Thulborn, 1996) were consistent with this prediction. MacDonald and Christiansen ignored this new information from brain imaging about the cognitive processing of sentences in the very task for which they offered a model.

The capacity theory also correctly predicted that the amount of brain activation associated with the use of a verbal strategy for sentence–picture comparison should be lower for high-span than for low-span subjects (Reichle, Carpenter, & Just, 2000). The reasoning is simply that people with a larger capacity have to utilize a smaller proportion of their capacity to process a given sentence. Thus, the capacity theory has had something useful to say both about sentence complexity effects and about individual differences. The continuous interaction between the predictions of the theory and the results of new fMRI studies guided by the theory has led to a successor architecture, Cortical Capacity-Constrained Collaborative Activation-based Production System (4CAPS) that retains the 3CAPS features but also provides an account of the amount and nature of brain activity in various cortical areas during cognition (Just, Carpenter, & Varma, 1999). A more detailed description of 4CAPS is beyond the scope of this article.

It may be possible to construct post hoc SRN-based accounts of such sentence-processing results on the basis of some yet-to-be-specified mapping between the MacDonald and Christiansen model and brain function, but MacDonald and Christiansen’s approach to date has been decidedly nongenerative on such issues. In particular, the single proposed link between the SRN and observable phenomena (such as word-reading times) is already tenuous; it is unclear how a second, independent link between the SRN and the amount of brain activity would be made. So although MacDonald and Christiansen claimed a linkage between their model and biological factors, their approach, in fact, has had little to say about any underlying biological mechanisms.

MacDonald and Christiansen also ignored the large number of studies in cognitive neuroscience establishing the neuroanatomic substrates of the WM systems associated with various domains other than sentence comprehension. These studies range from neuroimaging investigations of verbal and spatial WM tasks (e.g., Carpenter, Just, & Reichle, 2000; Cohen et al., 1997; Smith & Jonides, 1997) to single-unit studies with primates (e.g., Goldman-Rakic, 1987). In the face of mounting neuroscience evidence for the reality of WM and the increasing ability to characterize its components and its functional properties, MacDonald and Christiansen’s attempts to relabel it and define away its structural and functional reality are quixotic.

Limitations of MacDonald and Christiansen’s (2002) SRN Model of Sentence Comprehension

MacDonald and Christiansen proposed an SRN model, which they used to fit some aspects of the King and Just (1991) data. However, the SRN model is inadequate in a number of respects.

The SRN Model Does Not Compute Thematic Representations

The task of sentence comprehension is to transform word-level representations into thematic representations from which comprehension questions can be answered and discourse representations constructed. CC READER performs this task, taking perceptual encodings of words as input and producing thematic- (case-) role representations as output sufficient for answering *WH* (i.e., who, what, when, where, why, and how) questions and building propositions for discourse processing.

The model advanced by MacDonald and Christiansen does not construct a thematic representation. Rather, it produces a predicted distribution of grammatical classes for each successive word. However, predicting grammatical class distributions is not the raison d'être of sentence comprehension. Theories of sentence comprehension are expected to make contact with theories of word recognition, orthographic and phonological processing, and lexical access to specify the delivery of representations of an input stream of words, optionally annotated with syntactic and semantic information. Theories of discourse comprehension assume theories of sentence comprehension that specify thematic or propositional representations of sentences. A theory of sentence comprehension should span the gap between word and discourse theories. The SRN model fails to do this. It performs the wrong task. Ironically, it even fails to do the word recognition portion of sentence comprehension, despite MacDonald and Christiansen's claims of communion between models at the word and sentence level.

At best, the SRN renders a grammaticality judgment—in this case, that the sentence was syntactically well formed. This task may possibly bear on the processing of syntax, but it is not the task of participants in sentence comprehension studies, let alone the task of comprehension in the real world. Sentence comprehension requires that a thematic representation be computed, that understanding be achieved. On this, the SRN model is silent and thus incomplete.

The SRN Model Fits the King and Just (1991) Data Poorly

MacDonald and Christiansen claimed that the SRN model fits the King and Just data, but a closer look reveals a serious problem. Although the SRN model may learn grammatical sequences, it does not apply its knowledge in a psychologically realistic manner. This problem surfaces when the SRN model underpredicts reading time on the last word of the embedded clause, especially for the subject-relative sentence (MacDonald and Christiansen's Figure 2 and Just and Carpenter's Figure 9). Whereas both groups of human participants take longer on the clause-final word of subject relatives (the word *senator*) than on the average of the preceding four words, the SRN has a lower grammatical prediction error here. MacDonald and Christiansen (their Footnote 3) attempt to explain their model's misprediction in terms of a post hoc account related to variations in the length of the embedded clause. The SRN's failure of fit is particularly telling because it comes at a point in the sentence where there is an extra WM load. The observed phenomenon is predicted by the CC READER model. Center-embedded sentences require that when the embedded clause is encountered, the processing of the main clause be interrupted and the partial representation buffered while the embedded clause is processed. These additional storage demands consume some of the CC READER model's limited WM resources, slowing down its processing at the end of the embedded clause, particularly if the embedded clause increases in length. This slowing down provides a principled account of the longer processing times observed at the end of the embedded clause in the behavioral data. The SRN lacks any comparable mechanism to deal with a temporary high storage load encountered in the course of processing a sentence.

The SRN Model Lacks Empirical Scope

MacDonald and Christiansen fit their model to some of the results of one study (King & Just, 1991). They sketch in-principle accounts of other findings (e.g., MacDonald, Just, & Carpenter, 1992), but there is a large gap between a verbal account and a concrete model. CC READER and other 3CAPS models of sentence comprehension (Goldman & Varma, 1995; Haarmann, Just, & Carpenter, 1997; Thibadeau, Just & Carpenter, 1982; Varma & Goldman, 1996) are concrete models of a broad range of lexical, syntactic, thematic, discourse, and anaphoric phenomena. Moreover, MacDonald and Christiansen have carefully chosen which study to attempt to simulate, namely one whose results are driven almost exclusively by a structural phenomenon. Much additional work remains for MacDonald and Christiansen to do before they have demonstrated an adequate scope for the SRN model.

The SRN Model Is Structurally Similar to 3CAPS and CC READER

To the degree the MacDonald and Christiansen model is correct, it may be because it implicitly includes a structural analog to WM. What distinguishes SRNs from feed-forward connectionist networks is a layer of context units that allows the activation states of hidden units to feed back into processing with new input at the next time step. The context units store the partial products of current processing for use in future processing. These units are critical for the processing that integrates over inputs that are distributed over time. The context units of SRNs play a role that is a partial counterpart to the WM buffer in a production system, storing partial products and influencing future processing. Varying the number of context units would affect the SRN model's performance, just as varying the WM capacity of 3CAPS affects CC READER's performance.

SRNs May in Principle Be Incapable of Capturing Sentence Comprehension

Can connectionist networks adequately model sentence comprehension without simply reimplementing the main formalisms of symbolic accounts—discrete algorithms using rules and representations? Proponents of symbolic views have long argued that connectionist accounts may complement but cannot fully co-opt symbolic theories (e.g., Pinker, 1991). It is interesting that many connectionist theorists are now reaching the same conclusion, actively incorporating symbolic mechanisms in their models. For example, Smolensky (1999), according to Christiansen and Chater (1999), "suggests that progress requires a match between insights from the generative grammar approach in linguistics, and the computational properties of connectionist systems (e.g. constraint satisfaction)" (p. 15). Christiansen and Chater posed the rhetorical question of whether "promising initial results can be scaled up to deal with the complexities of real language, or whether a purely connectionist approach is beset by fundamental limitations, so that connectionism can only succeed by providing reimplementations of symbolic methods" (p. 10). This question remains to be answered.

3CAPS as Both a Connectionist and a Symbolic Architecture

MacDonald and Christiansen (2002) questioned whether 3CAPS is a hybrid architecture, claiming that it is symbolic but not connectionist. They attempted to guard the connectionist tradition, not only against 3CAPS incursions but also against more conventional connectionist modelers who choose to explore the unorthodox. ("Of course, it is possible to construct a hybrid model which includes a separate working memory, . . . but this approach violates the predominant connectionist approach to language" [MacDonald & Christiansen, 2002, Footnote 10, p. 49]). 3CAPS is neither an orthodox symbolic nor an orthodox connectionist architecture. Rather, it combines symbolic and connectionist mechanisms to realize a particular view of cognition: the collaboration of multiple knowledge sources in a capacity-constrained environment to perform a task. It is in this larger sense that 3CAPS is a hybrid architecture, its four architectural principles specifying its full name: a capacity-constrained collaborative activation–based production system architecture. The section below describes the component themes of the hybrid 3CAPS and points out several common ancestors with modern connectionism.

3CAPS Is a Production System

The symbolic endowment of 3CAPS traces back to Newell's (1973) advocacy of production systems as a cognitive architecture. A production system encodes knowledge procedurally in terms of productions that require no central executive because of their self-scheduling capability. Productions possess variable-binding capabilities extremely useful for modeling high-level cognition, such as the ability to concisely encode relations among classes of elements without having to list each possible instance of the relation. For example, in the case of sentence processing, variable binding makes it possible for a single production to specify how determiners and nouns combine to form noun phrases without having to specify it for every possible determiner–noun pair. The combinatorial and variable-laden nature of WM elements allows productions to naturally express complex structured representations, such as linguistic representations and problem-solving strategies.

Conventional symbolic architectures have shortcomings as well. They are brittle—the addition or subtraction of a seemingly small piece of knowledge can grind cognition to a halt—and they are often serial in nature, which is problematic because some aspects of cognition naturally suggest parallel processing.

3CAPS Is Activation Based

The connectionist aspects of 3CAPS stem from work on spreading activation models of semantic memory in the late 1960s and early 1970s (Anderson & Bower, 1973; Collins & Loftus, 1975; Quillian, 1968). Spreading activation models influenced the localist connectionist networks of early modern connectionism—for example, the interactive activation model of word recognition (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). 3CAPS can be construed as a localist connectionist network. The elements of WM make up the nodes of the network. Each element is annotated with an activation value. Productions implement links between nodes. The primary action of productions is to direct activation from one element to another, modulo some weight. On each cycle of processing, activation is directed to elements across weighted links defined by productions firing in parallel. Moreover, the productions fire reiteratively over successive cycles, accomplishing a graded form of processing that can match a human temporal performance profile. Appendix A formally describes the activation dynamics in 3CAPS.

3CAPS Is a Symbolic–Connectionist Hybrid

The properties of connectionist processing in 3CAPS mitigate the disadvantages associated with its symbolic side. It is a naturally parallel system that can bring large amounts of knowledge to bear to solve problems. The graded nature of its symbolic representations softens their otherwise brittle nature, granting 3CAPS some robustness in the face of changes to its knowledge. At the same time, 3CAPS retains the capabilities inherited through its symbolic ancestry, variable binding and combinatorial structures, in particular—mechanisms that have resisted a satisfying connectionist implementation. The complementary properties of its two processing styles constitute the hybrid nature of 3CAPS. Of course, 3CAPS is not the only architecture that has merged symbolic and connectionist mechanisms to enduring effect. Others include ACT* (Anderson, 1983), the "Induction" framework (Holland, Holyoak, Nisbett, & Thagard, 1986), and XAPS (Rosenbloom & Newell, 1987), the progenitor of Soar (Newell, 1990). These three architectures explicitly list the first generation of connectionist networks—parallel associative models of memory (Hinton & Anderson, 1981) and the word-reading model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982)—as influences. In the early and mid 1980s, before the boundaries between the symbolic and connectionist camps had been so strongly drawn, it was natural for Rumelhart and McClelland (1986, p. 549) to notice the family resemblance between ACT*, CAPS, and connectionist models.

3CAPS Embodies Collaborative Processing

3CAPS benefited from other modelers' lessons that cognition is a collaborative affair. The collaborative processing style was pioneered by the Hearsay-II system, which attempted in the 1970s to develop a real-time speech recognition and understanding system with a 1,000-word vocabulary (Erman, Hayes-Roth, Lesser, & Reddy, 1980; Reddy, Erman, Fennel, & Neely, 1973). Hearsay-II was composed of a number of knowledge sources, each expert at a particular level of language processing. Moreover, the different knowledge sources were not modular but collaborated interactively, revising and constraining each other's processing. The knowledge sources collaborated by means of a shared structure called the *blackboard*. The metaphor is of several colleagues with expertise in different domains cooperatively solving a problem in front of a blackboard, each modifying the work of the others, a solution emerging from their collective action.

Hearsay-II's casting of cognition as collaborative processing profoundly influenced subsequent architectures. It shaped the processing style embodied in CAPS: a "collaborative execution of processes operating at different levels of analysis" (Thibadeau et al., 1982, p. 157). It is a central reference of the CAPS article, the

XAPS article (Rosenbloom & Newell, 1987), the ACT* book (Anderson, 1983), and the "Induction" book (Holland et al., 1986). It lurks behind the more contemporary hybrid Copycat and Jumbo programs of Hofstadter and colleagues to such a degree that Hofstadter remarks that "the influence of the Hearsay II project in speech understanding on my work cannot be overstated" (Hofstadter & the Fluid Analogies Research Group, 1995, p. 91).

Hearsay-II also influenced connectionists of the time. McClelland, Rumelhart, and Hinton (1986) claimed that the key difference between conventional and parallel distributed processing is sensitivity to "multiple simultaneous constraints" (p. 4), citing the interaction of syntactic and semantic processing as an example (pp. 6–7). They also claimed that Hearsay-II "inspired" the word-reading model (McClelland et al., 1986, p. 43) and "the programmable blackboard model of reading" (McClelland, 1986, p. 122). When confronted with the question of what led him from serial processing to connectionism, Rumelhart (in an interview conducted by Baumgartner & Payr, 1995) offered this:

> In my own mind, I had been intrigued by parallel processing ideas for a long time. I was very much impressed by the work of Reddy and his colleagues at Carnegie-Mellon when they had developed the so-called HEARSAY speech recognition system, where the idea of cooperative computation was central . . . . These ideas haunted me for really a long time. (p. 191)

Collaborative processing, the processing style of 3CAPS, is a major element of its heritage shared with connectionist architectures.

3CAPS Is Capacity Constrained

The main change from CAPS to 3CAPS was the proposal of a capacity constraint to reflect the intrinsic limitations of the human cognitive architecture that shape information processing. Although CAPS had construed information processing as a resource-consuming activity, the assumption had little impact as long as the resource was unlimited. 3CAPS added an explicit constraint on the operational capacity of WM, affecting processing and storage conjointly. The capacity constraint is formalized in Appendix B. Because cognition takes place in a constrained arena, representations and processing contend for scarce resources. The capacity constraint introduces a competitive element to the collaborative processing style of 3CAPS because alternative representations vie for limited activation, in some ways analogous to the lateral inhibition mechanisms in connectionist models.

The Role of Experience and Learning in 3CAPS

The capacity theory easily allows a role for experience in explaining individual differences in WM, contrary to MacDonald and Christiansen's speculations. Appendix C derives a learning rule that tunes the weights of 3CAPS productions with experience. The weight of a production governs the rate at which it propagates activation. The weights are increased with practice, as shown in Figure 1. This produces a decrease in the amount of time the reiterative firing of a production will take to perform a given function, following the well-documented power law of practice (Newell & Rosenbloom, 1981). Further note that another factor influences the time for a production to achieve its goal: the capacity of WM. (This is described in the next section and in Appendix C.) Thus, 3CAPS illuminates the relationship between experience and processing speed within the capacity theory.

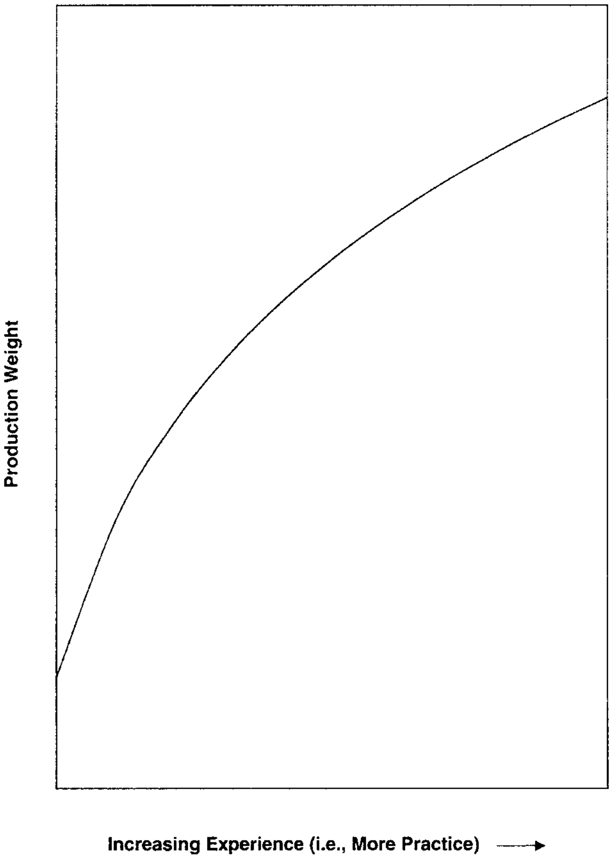

Figure 1. The weights of productions increase with experience and practice.

In summary, the architectural principles behind 3CAPS combine the strengths of symbolic and connectionist processing styles while avoiding their respective weaknesses. Its computational mechanisms are harnessed to deliver collaborative processing in a capacity-constrained WM.

3CAPS Accounts of WM Phenomena

We discuss here two of the phenomena to which MacDonald and Christiansen (2002) alluded, indicating how the 3CAPS formalism provides a broad integrating framework that coherently accounts for these diverse phenomena.

The Capacity Theory Account for Long-Term WM (LT-WM) Phenomena

According to LT-WM theory (Ericsson & Kintsch, 1995), people have short-term working memories (ST-WM) of similar capacities. With practice, according to Ericsson and Kintsch's (1995) view, experts develop hierarchical retrieval structures, systems of distinctive retrieval cues that enable them to store information in and retrieve information from long-term declarative memory

(LTDM) with minimal interference. The portion of LTDM accessed by retrieval structures is termed LT-WM; it augments short-term WM in offering fast and reliable storage and access. Perhaps the best evidence for LT-WM is the case of S.F., who acquired a span in excess of 80 digits after months of practice (Chase & Ericsson, 1981). That he utilized hierarchical retrieval structures in LT-WM to accomplish this feat was inferred from the timing of his utterances during recall.

LT-WM is a verbally stated theory that makes qualitative predictions. It has been formally instantiated in a 3CAPS model of S.F. (Varma, 1996). The model contains three sets of productions. The first set implements an LTDM, specifically the search of associative memory theory (Raaijmakers & Shiffrin, 1981). The second set operates during encoding. It annotates digits being stored in LTDM with distinctive cues that form an implicit hierarchical structure. The third set of productions implements retrieval. With knowledge of the cueing scheme, it traverses the retrieval structure, retrieving the to-be-recalled digits. The model captures the important behavioral regularities of the phenomenon. In particular, it simulates the timing of utterances during recall, providing striking evidence for the underlying hierarchical retrieval structures. The 3CAPS model is an existence proof of the consistency of the capacity and LT-WM theories.

MacDonald and Christiansen's claim that their experiential approach and the LT-WM theory are consistent is true at the level that both predict increased domain-specific WM performance with practice. However, for the similarity to apply at a deeper level, MacDonald and Christiansen would have to develop an SRN implementation of LTDM and show that it learns a counterpart of hierarchical retrieval structures with training. The SRN would then have to encode and recall an 80-digit sequence in order, with no interference, while matching the recall cadence reported by Chase and Ericsson (1981). Until such an existence proof is offered for the SRN, it is unclear whether MacDonald and Christiansen can account for the key LT-WM phenomena.

Effects of Aging on Comprehension

MacDonald and Christiansen questioned the capacity theory's attribution of age-related decrements in comprehension to WM decrements. They cited Salthouse's (1996) attribution of aging effects to the speed of processing. Salthouse's psychometric analyses, which identified speed as the factor that changes with age, did not focus on experimental studies of sentence comprehension in the elderly. By contrast, Small, Kemper, and Lyons (1997) specifically examined the basis of sentence comprehension decrements in Alzheimer's disease by manipulating the rate of stimulus presentation. Small et al. concluded from their results that "working memory capacity (Just & Carpenter, 1992) plays a more significant role than does a general speed factor (Salthouse, 1996) in sentence comprehension" (p. 9). Thus, the capacity theory remains one of the viable candidates for explaining age-related changes in sentence comprehension, although the empirical issues are far from settled.

MacDonald and Christiansen presented a very general account of how their approach might apply to age-related changes in comprehension, reiterating that the interaction between experience and biological factors is likely to be complicated. Only in their Footnote 9 (p. 48) did MacDonald and Christiansen begin to enumerate some of the myriad complexities with which they have not dealt. These include hypotheses about the effects of aging and experience predicting opposite effects in the elderly, different kinds of language experience being differentially helpful for performance in a psycholinguistic experiment, and the effects of experience producing diminishing returns with aging. Therefore, despite MacDonald and Christiansen's claims that their account is consistent in a general way with other theories of cognitive aging, the detail of the application of their approach to aging effects on the speed and accuracy of sentence comprehension is absent.

It should be noted that a decreased capacity account of age-related decline entails a processing-speed decline under some circumstances. A typical 3CAPS production has the goal of directing activation from a source element to a target element, modulo some weight. It will fire iteratively until the activation of the target exceeds some threshold, resembling a test-operate-test-exit unit (Miller, Galantner, & Pribram, 1960). Appendix C states the inverse relationship between the expected time for a production to accomplish its goal and the capacity of WM. Thus, if WM capacity decreases, as might occur with aging, the expected completion time increases and hence processing speed decreases. Figure 2 depicts the relationship given in Appendix C.

Conclusion

Although MacDonald and Christiansen (2002) have generated a post hoc account of some aspects of the King and Just (1991)

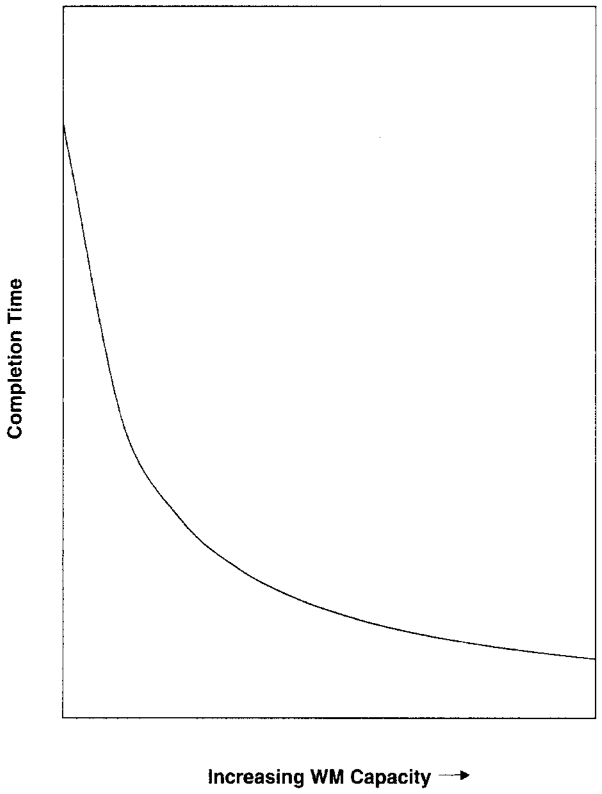

Figure 2. Productions achieve their goals more quickly the larger the capacity of working memory (WM).

results, it contributes little new understanding, fails to generate significant new questions or answers, and completely fails to address a near-decade's new understanding of human language function. Instead of providing a road map to the future, it attempts a minor revision of the past. MacDonald and Christiansen ask "Where is working memory? To the extent that it is useful to talk about working memory within these systems, it is the network itself" (p. 38). We completely agree that usefulness is the relevant criterion. Insofar as MacDonald and Christiansen's characterization will prove in the future to be useful in some way, then they have made a contribution. To the extent that the attempt to do away with WM is just an exercise in formalism and semantics, then there is not much contribution made and it is much ado about nothing.

However, WM for language, contrary to MacDonald and Christiansen's perspective, is more than just a network. It is the coordinated activity of a dynamic, resource-consuming system with an architecture and with resource constraints. Advances in functional neuroimaging methods have made it possible to identify the cortical substrate of WM for sentence comprehension, to measure the activation in a set of cortical areas during comprehension tasks, and to manipulate the amount and the precise locus of the activation by varying the comprehension conditions. The existence proof provided by a simple recurrent network model does not begin to approach an account for such phenomena. The capacity theory, by contrast, provides a conceptual framework that is proving to be increasingly useful in a number of research areas, and its vitality is demonstrated by its evolution and encompassing of new domains.

References

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

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory.* Washington, DC: Winston.

Baumgartner, P., & Payr, S. (1995). *Speaking minds: Interviews with twenty eminent cognitive scientists.* Princeton, NJ: Princeton University Press.

Callicott, J. H., Mattay, V. S., Bertolino, A., Finn, K., Coppola, R., Frank, J. A., et al. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cerebral Cortex, 9,* 20–26.

Caplan, D., & Waters, G. (2002). Working memory and connectionist models of parsing: A response to MacDonald and Christiansen (2002). *Psychological Review, 109,* 66–74.

Card, S. K., Moran, T. P., & Newell A. (1983). *The psychology of human–computer interaction.* Hillsdale, NJ: Erlbaum.

Carpenter, P. A., Just, M. A., & Reichle, E. D. (2000). Working memory and executive function: Evidence from neuroimaging. *Current Opinion in Neurobiology, 10,* 195–199.

Chase, W. G., & Ericsson, K. A. (1981). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning* (Vol. 16, pp. 1–58). New York: Academic Press.

Christiansen, M. H., & Chater, N. (1999). Connectionist natural language processing: The state of the art. *Cognitive Science, 23,* 417–437.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature, 386,* 604–608.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82,* 407–428.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3,* 422–433.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211–245.

Erman, L. D., Hayes-Roth, F., Lesser, V. R., & Reddy, D. R. (1980). The Hearsay-II speech-understanding system: Integrating knowledge to resolve uncertainty. *ACM Computing Surveys, 12,* 213–253.

Gibson, E., & Schutze, C. T. (1999). Disambiguation preferences in noun phrase conjunction do not mirror corpus frequency. *Journal of Memory and Language, 40,* 263–279.

Goldman, S. R., & Varma, S. (1995). CAPping the construction–integration model of discourse comprehension. In C. A. Weaver, S. Monners, & C. R. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 337–358). Hillsdale, NJ: Erlbaum.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.), *Handbook of Physiology—The Nervous System* (Vol. 5, pp. 373–417). Bethesda, MD: American Physiological Society.

Goldstein, G., Minshew, N., & Siegel, D. (1994). Age differences in academic achievement in high-functional autistic individuals. *Journal of Clinical and Experimental Neuropsychology, 16,* 671–680.

Haarmann, H. J., Just, M. A., & Carpenter, P. A. (1997). Aphasic sentence comprehension as a resource deficit: A computational approach. *Brain and Language, 59,* 76–120.

Hinton, G. E., & Anderson, J. A. (1981). *Parallel models of associative memory.* Hillsdale, NJ: Erlbaum.

Hofstadter, D., & the Fluid Analogies Research Group. (1995). *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought.* New York: Basic Books.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning, and discovery.* Cambridge, MA: MIT Press.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122–149.

Just, M. A., Carpenter, P. A., & Keller, T. A. (1996). The capacity theory of comprehension: New frontiers of evidence and arguments. *Psychological Review, 103,* 773–780.

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996, October 4). Brain activation modulated by sentence comprehension. *Science, 274,* 114–116.

Just, M. A., Carpenter, P. A., & Varma, S. (1999). Computational modeling of high-level cognition and brain function. *Human Brain Mapping, 8,* 128–136.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30,* 580–602.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: A comment on Just & Carpenter (1992) and Waters & Caplan (1996). *Psychological Review, 109,* 35–54.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology, 24,* 56–98.

McClelland, J. L. (1986). The programmable blackboard model of reading. In J. L. McClelland, D. E. Rumelhart, & The PDP Research Group (Eds.), *Parallel distributed processing: Vol. 2. Psychological and biological models* (pp. 122– 169). Cambridge, MA: MIT Press.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

McClelland, J. L., Rumelhart, D. E., & Hinton, G. E. (1986). The appeal of parallel distributed processing. In D. E. Rumelhart, J. L. McClelland, & The PDP Research Group (Eds.), *Parallel distributed processing: Vol 1. Foundations* (pp. 3–44). Cambridge, MA: MIT Press.

Miller, G. A., Galantner, E., & Pribram, K. H. (1960). *Plans and the structure of behavior.* New York: Holt, Rinehart & Winston.

Newell, A. (1973). Production systems: Models of control structures. In W. G. Chase (Ed.), *Visual information processing* (pp. 463–526). New York: Academic Press.

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

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill aquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their aquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.

Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 141–183). Hillsdale, NJ: Erlbaum.

Pinker, S. (1991). Rules of language. *Science, 253,* 530–535.

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

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93–134.

Reddy, D. R., Erman, L. D., Fennel, R. D., & Neely, R. B. (1973). The Hearsay speech understanding system: An example of the recognition process. *Proceedings of the Third International Joint Conference on Artificial Intelligence,* 185–193.

Reichle, E. D., Carpenter, P. A., & Just, M. A. (2000). The neural bases of strategy and skill in sentence-picture verification. *Cognitive Psychology, 40,* 261–295.

Roberts, R., & Gibson, E. (in press). Individual differences in sentence memory. *Journal of Psycholinguistic Research.*

Rosenbloom, P. S., & Newell, A. (1987). Learning by chunking: A production-system model of practice. In D. Klahr, P. Langley, & R. Neches (Eds.), *Production system models of learning and development* (pp. 221–286). Cambridge, MA: MIT Press.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60–94.

Rumelhart, D. E., & McClelland, J. L. (1986). Future directions. In J. L. McClelland, D. E. Rumelhart, & The PDP Research Group (Eds.), *Parallel distributed processing: Vol. 2. Psychological and biological models* (pp. 547–552). Cambridge, MA: MIT Press.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428.

Small, J. A., Kemper, S., & Lyons, K. (1997). Sentence comprehension in Alzheimer's disease: Effects of grammatical complexity, speech rate, and repetition. *Psychology and Aging, 12,* 3–11.

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology, 33,* 5–42.

Smolensky, P. (1999). Grammar-based connectionist approaches to language. *Cognitive Science, 23,* 589–613.

Thibadeau, R., Just, M. A., & Carpenter, P. A. (1982). A model of the time course and content of reading. *Cognitive Science, 6,* 157–203.

Varma, S. (1996). *A computational model of long-term working memory and its application to skilled memory and discourse comprehension* [Technical report]. Nashville, TN: Learning Technology Center, Vanderbilt University.

Varma, S., & Goldman, S. R. (1996, January). *Resonance-based reinstatement of anaphoric antecedents.* Poster presented at the Seventh Annual Winter Text Conference, Jackson Hole, WY.

Waters, G., & Caplan, D. (1996). The capacity theory of sentence comprehension: A critique of Just and Carpenter (1992). *Psychological Review, 103,* 761–772.

Appendix A

Activation Dynamics of Capacity-Constrained Collaborative Activation–Based Production System (3CAPS)

3CAPS can be construed as a localist connectionist architecture under the following mapping.

Time is discrete and comes in units called cycles. Consider the sequence of cycles $T = < 0, 1, \ldots >$.

Construe the elements of working memory as nodes in a localist network; denote the *i*th element as E_i. Each element has an activation that reflects (roughly) the degree to which it is believed important for current (and future) processing. Denote the activation of element E_i at time $t \in T$ as $A_i(t)$. This value is partly a function of the element's activation on the previous cycle and partly a function of its net input for the current cycle. The former quantity, called the storage demand, is denoted as $S_i(t)$ and defined as

$$S_i(t) = A_i(t - 1). \qquad \text{(A1)}$$

The latter quantity is called the processing demand and denoted as $P_i(t)$. If $WEIGHT_{ij}(t)$ denotes the dynamic weight of the connection from source element E_j to target element E_i at time t, then the processing demand of (i.e., net input to) E_i can be defined as

$$P_i(t) = \Sigma_j\, A_j(t - 1) \times WEIGHT_{ij}(t). \qquad \text{(A2)}$$

The connection weights between elements are defined dynamically. On each cycle, a number of productions will be instantiated and fired in parallel. Each will direct activation from a source to a target element modulo the weight associated with the production. $WEIGHT_{ij}(t)$ is the sum of the production weights that direct activation from E_j to E_i at time t. The raw activation demanded by element E_i at time t is denoted $R_i(t)$ and defined as

$$R_i(t) = S_i(t) + P_i(t). \qquad \text{(A3)}$$

Elements are granted their raw activation demands:

$$A_i(t) = R_i(t). \qquad \text{(A4)}$$

Note, additionally, that the dynamically varying connection weights allow 3CAPS to transcend its otherwise linear activation dynamics.

Appendix B

Capacity-Constrained Working Memory

This appendix extends the activation dynamics described in Appendix A to include a capacity-constrained working memory.

The total raw activation demand on 3CAPS at time t is denoted $R(t)$ and is defined as $R(t) = \Sigma_j R_j(t)$. The maximum activation capacity of 3CAPS is denoted by *CAP*. If there is sufficient capacity to meet the demand, that is, $R(t) \leq CAP$, then the demand is satisfied; that is, all elements are allocated the activation they demand. However, if the available capacity is insufficient to meet the demand, that is, $R(t) > CAP$, then the activation request of each element is scaled back proportionately. We can subsume both of these cases by defining a scaling factor $d(t)$ as

$$d(t) = \min\{1, CAP/R(t)\}. \qquad \text{(B1)}$$

The activation of element E_i at time t, previously defined by Equation A4, is now

$$A_i(t) = d(t) \times R_i(t). \qquad \text{(B2)}$$

Note that the scaling factor $d(t)$, which is a dynamic function of the total raw activation demand $R(t)$ and the available activation *CAP*, is another nonlinear component of the activation dynamics of 3CAPS.

Appendix C

Learning Rule for Capacity-Constrained Activation–Based Production System (3CAPS)

This appendix derives a learning rule for 3CAPS that decreases the time required for a production to achieve its goal according to the power law of practice. It works by modifying the weights on productions and therefore modifying the rates at which they propagate activation.

The typical 3CAPS production has the following form:

Production k:
IF a pattern of elements exists, including a source element E_j, and the target element E_i has activation $A_i\ (t-1) < THRESH_k$
THEN direct $A_j(t-1) \times WEIGHT_k$ units of activation to E_i.

This production, indexed by k, has the goal of directing activation from a source element E_j to a target element E_i, modulo the weight $WEIGHT_k$. It will fire iteratively over time (cycles) until the activation of E_i exceeds the threshold $THRESH_k$. ($WEIGHT_k$ and $THRESH_k$ are parameters of the production.)

The time required for a production to achieve its goal is estimated before deriving a learning rule. Assume that the total activation demanded during the interval T is relatively constant:

$$d(t) \approx d \text{ for } t \in T. \qquad \text{(C1)}$$

Note that when $R(t) > CAP$, d is proportional to the activation capacity:

$$d \propto CAP \text{ by Equations B1 and C1.} \qquad \text{(C2)}$$

It can be shown (with a few additional assumptions) that the approximate time required for production k to achieve its goal is

$$t_k \approx [THRESH_k - A_i(0)]/[d \times A_j(0) \times WEIGHT_k]. \qquad \text{(C3)}$$

(Space considerations preclude including the derivation here.) This equation makes intuitive sense; the time is the ratio of the amount of activation the goal requires, $THRESH_k - A_i(0)$, and the amount of activation directed per unit time, $d \times A_j(0) \times WEIGHT_k$.

Note that because the time required for a production to achieve its goal is inversely proportional to d (i.e., Equation C3), and because d is proportional to *CAP* when the total demand on activation exceeds the available capacity (i.e., Equation C2), then the time for a production to achieve its goal is inversely proportional to the capacity of working memory.

To derive a learning rule, denote the time for a production k to achieve its goal the nth time as $t_k^{(n)}$. The power law of practice states that this time follows a power function:

$$t_k^{(n)} = t_k^{(l)} \times n^{-a}. \qquad \text{(C4)}$$

(Note that the constant a is typically in the interval [.2, .6]; Card, Moran, & Newell, 1983.) To derive a learning rule that tunes productions in such a way that their processing time decreases according to Equation C4, first, derive a learning factor $L^{(n)}$ that relates the time required for production k the nth and $(n-1)$th times:

$$t_k^{(n)} = L^{(n)} \times t_k^{(n-1)}. \qquad \text{(C5)}$$

Solving for $L^{(n)}$ produces

$$\begin{aligned} L^{(n)} &= t_k^{(n)}/t_k^{(n-1)} = (t_k^{(1)} \times n^{-a})/(t_k^{(1)} \times (n-1)^{-a}) \\ &= [n/(n-1)]^{-a}. \qquad \text{(C6)} \end{aligned}$$

Substituting Equations C3 and C6 into C5 yields

$$\begin{aligned} t_k^{(n)} &\approx [n/(n-1]^{-a} \\ &\times [THRESH_k^{(n-1)} - A_i(0)]/[d \times A_j(0) \times WEIGHT_k^{(n-1)}]. \qquad \text{(C7)} \end{aligned}$$

Of the two production parameters $THRESH_k^{(n-1)}$ and $WEIGHT_k^{(n-1)}$, the second one directly supports a learning rule:

$$WEIGHT_k^{(n)} = [n/(n-1)]^{a} \times WEIGHT_k^{(n-1)}. \qquad \text{(C8)}$$

Thus, the learning rule increases the weight of the production with experience and practice. This increases the amount of activation the production directs to its target element per unit time, decreasing the time required to activate the target element above threshold—that is, for the production to achieve its goal. This decrease in time follows the power law of practice.

Received July 3, 2000
Revision received December 5, 2000
Accepted December 5, 2000 ■