Developmental Science 6:4 (2003), pp 440–447

Comparisons of developmental modeling frameworks and levels of analysis in cognition: connectionist and dynamic systems theories deserve attention, but don't yet explain attention

Nelson Cowan

Department of Psychological Sciences, University of Missouri, USA

Abstract

In this article, an investigator who is outside of the connectionist / dynamic systems (CD) tradition, but is sympathetic to many of its theoretical stances, considers whether these are separate approaches and what the two approaches mean for the development of information processing. It is argued that these approaches have not been defined in the target articles of this special issue narrowly enough to make them clearly separate approaches. Rather, they appear to emphasize different aspects of a common approach. It is also suggested that they do not form an overarching theory of development but do form an important theoretical approach at one level of analysis. In particular, one may sometimes, but not always, have to get into nonlinear interactions between faculties of the mind to predict and explain behavior. Finally, an attempt is made to challenge the CD approach to become more global by confronting the issue of how attention and conscious awareness should be represented in these models, and how the development of them may occur. Evidence of changes in automatic and attention-related parameters of processing may provide a scaffolding upon which better CD models can be constructed.

Introduction

This article contributes to the special issue from the point of view of an 'outsider'. After considering my qualifications for that job, I will briefly address the dominant theme of whether the connectionist and dynamic systems theoretical frameworks are the same or different. Then I will make some comments regarding my beliefs as to whether the two of them, viewed together as a connectionist / dynamic systems (CD) framework, may form an overarching theory, and the related question of how and when it is good to use the CD framework. Last, I will discuss a domain of human behavior that may pose one of the greatest challenges if the CD framework is to achieve the greatest breadth: the domain of attention, touching on conscious thought, free will, working memory, executive control and serial processing.

The author's theoretical vantage point

In what way am I an outsider? The purpose of the special issue is to compare connectionist and dynamic systems theories of development. I have never worked on a connectionist or dynamic systems model and rarely engage in computer or mathematical modeling (although I program at times and have done a little modeling). I have written about development (e.g. Cowan, Elliott & Saults, 2002) and have written about overarching models of human information processing (e.g. Cowan, 1988, 1995, 1999). However, the developmental work has been aimed at clarifying certain elementary mechanisms (processing speed; processing capacity; memory persistence) and the overarching model is meant as a description of what appears to be *necessary* to include in the human information processing system, with many important questions deliberately left unresolved or undecided in my theoretical framework and in its graphic representation.

I am not the furthest outsider in that I tend to agree with some of the basic premises that underlie the CD framework. That framework is one in which specialized functions can be seen to emerge from lower-level neural activity so that learning plays a large role. For example, the framework does not place great stock in the Chomskian notion that innate, specialized mechanisms determine everything interesting in language and its development. I have disagreed with the Chomskian type of stance since my graduate work in the 1970s. Importantly, this stance leaves unresolved the issue of how the

Address for correspondence: Nelson Cowan, Department of Psychological Sciences, University of Missouri, 18 McAlester Hall, Columbia, MO 65211 USA; e-mail: CowanN@missouri.edu

brain would be able to send each incoming stimulus to its proper module. Instead, presumably like CD theorists, I favor the notion that language comprehension and production function more like problems to be solved in other domains. Thus, I am a 'near outsider' rather than a far one. However, I am a strong believer in the notion that the brain operates through a dual system that includes not only parallel processes, which seem to be naturally suited to connectionist approaches, but also serial processes under deliberate control, which may be adapted to the CD approach but perhaps do not emerge naturally from all such approaches. I also question the role of the embodiment of cognition that is espoused within some dynamic systems views, for reasons that I will explain.

A brief comparison of dynamic systems and connectionism

The theme of this special issue is whether connectionist and dynamic systems models are the same or different. That is difficult to answer with certainty when both approaches are in the process of growth and change. For example, Thelen and Bates (this issue) noted that in the earlier draft of their article, Bates 'placed a brisk "NO" for dynamic systems in the category "emphasis on mental representations", compared with a clear "YES" for connectionism'. 'In the current version of Table 1', they continued, 'we have changed that notion to read "NO until after 1994" . . .' This changing theoretical landscape was a central challenge for me as I tried to synthesize across the target articles.

I do not doubt the opinion, formalized in Thelen and Bates' Table 1 comparing theories, that the two approaches now differ primarily in that the dynamical systems approach is more theoretically committed to the notion of embodied cognition. The two views are similar in their emphasis that the interaction of lower-level units in a nonlinear manner can result in the emergence of aspects of cognition that one might not expect a priori, and that this emergence can occur across development. Prime examples include the production of overregularized past tense verbs, for connectionism, and the emergence and re-emergence of A-not-B errors with different task demands, for dynamic systems.

More generally, however, I doubt if the question of whether connectionism or dynamic systems are the same or different will come through as the most central or interesting question to typical readers (i.e. outsiders). The kinds of questions they are likely to ask are: 'What are connectionist and dynamic systems models like? What are their theoretical underpinnings? Where can they lead us? How do they help me to explain, and even predict, behavior? What time commitments would I have to make to use them, and how then could I use them?' The articles do address many of these questions along the way. Readers and researchers are likely to take what is best from each of these two seemingly compatible approaches and synthesize a new, combined approach from them. Thus, rather than focusing on the detailed comparisons of each approach, I opted to focus on what I perceive to be the important, broad issues from an outsider's perspective. These broad issues form the focus of the sections that follow.

CD as an overarching theory?

Thelen and Bates (this issue) suggested in their abstract that 'connectionism and dynamic systems theory are strong contenders for a general theory of development that holds true whatever the content domain'. Although I am enthusiastic about what CD theories have achieved recently, I do not see them as forming exactly an 'overarching' theory, for several interrelated reasons: the continued usefulness of the symbolic level of analysis, the issue of how much to trust a modeler's results, and doubts about the role of the embodiment principle found in dynamic systems theories.

First, whereas CD theories tend to work on a subsymbolic level of analysis in which basic processing elements are allowed to interact and behavior emerges, important theoretical work still might be accomplished at the symbolic level of representation. This is true even with the assumption that CD theories are not antithetical to the formation of mental representations. Symbolic representations can be useful even if they oversimplify the brain state. For example, working from a neo-Piagetian point of view, Halford, Wilson and Phillips (1998) could predict children's performance on a large variety of tasks by taking into account the number of stimulus dimensions that had to be combined (i.e. the relational complexity) and the child's processing capacity in that regard. Tasks included, for example, transitive inferences, conservation tasks and the comprehension of embedded sentences. Halford *et al.* did not have to state their hypotheses in a CD-based form in order to reach these conclusions. Thus, I see no clear sense in which the dynamic systems level, as opposed to the symbolic level, is 'overarching'. In my view, these are two different levels of analysis that are probably compatible and could be woven into a coherent story in the future.

In order to clarify this point, consider an imaginary conversation between Esther Thelen and Jean Piaget. (I take this liberty given the present, non-interactive venue.) Thelen could point out that the A-not-B error cannot be viewed as a case in which the child either has the object concept or not, given that the child's sensory and motor predicaments act as force fields that make the child seem either more or less sophisticated conceptually. I believe that Piaget would not be terribly surprised or concerned by this statement or by the new A-not-B results. He saw childhood and human history as two domains that can help to illustrate stages of knowledge structure that form a *logically necessary* progression. He knew that there are processing factors that influence the results, which led to his conception of progression that depends on the materials (horizontal décalage) and range of maturity (vertical décalage). However, for better or worse, he was not very interested in such processing factors. He probably would argue that task circumstances, such as whether the object is partly visible and whether the subject is standing or sitting, are influential mostly in the early stages of the acquisition of the concept and that there comes a time, after the concept is firmly held and completely acquired, after which they exert little influence. Thus, after the concept is attained, according to Piaget, one could make predictions based on the conceptual (or competence) level and no longer necessarily at the dynamic (or performance) level. At least, he would point to unfettered circumstances in which older children would show rule-based actions.

To this, a CD theorist might respond that the CD level is nevertheless overarching: 'If my computer model responds just like a human, at all levels of development, then why not simply use the computer model and abandon the attempt to think at a symbolic level?' In response to this, I would emphasize that a model, no matter how complete, is not itself a theory. Suppose that a modeler succeeds in writing a very detailed program that acts like a human being. Nevertheless, the programmer does not necessarily have a complete understanding of why the program responds this way. In an argument to absurdity, although many people actually create and grow a human being through conception and child-rearing, their creative success in no way implies that they understand how the child operates; they just know how to build one. Of course, the three target articles make it clear that considerable progress has been made not just in simulating behavior but also in understanding why CD models work. Still, the CD approach may not constitute a one-size-fits-all theory.

To summarize, the target articles have shown that a CD approach is *sufficient* in many circumstances. It remains to be shown that a CD approach is always *necessary*. Its strength has been largely in showing what is possible: how to get stages from continuous change, how to get what appear to be modular systems from undifferentiated nets plus input, how to get variability from one situation to the next, how to get U-shaped growth, and so on. One always must question, as did Massaro (1988), whether one needs recursive and nonlinear systems to account for all of this. Sometimes, the answer may be 'no'.

The second reason why I do not see CD theories as overarching stems from a healthy skepticism on my part: in order to use and understand CD models, most researchers must take a lot on faith. Researchers are consumers of others' research products who must trust others to a certain extent in order to preserve enough time to think for themselves when it counts. For example, we all must, to some extent, trust that the results that we read were as described. Most of us also trust that statistical programs will yield reasonable results. However, we do not want to trust too much. We are often unwilling to trust that the author of an article used an adequate research design if that design was not explicitly described, particularly if the research is in our area of expertise. Yet, in the case of CD modeling, we are challenged to trust aspects of computer modeling that most of us cannot verify. Consequently, if investigators from outside of the CD approach can find theoretical approaches that allow them to reach valid conclusions without considering the nonlinear interactions of low-level units, they are likely to capitalize upon such approaches.

I have learned to take what modelers tell me guardedly. Although many modeling results in psychology have been impressive, it also is true that two models based on incompatible assumptions both can seem successful (e.g. models of processing assuming a single memory system versus separate declarative and procedural systems; see Cowan, 1995, Chapter 1). Modelers become quite invested in their models and usually can account for most new data, if not through the central assumptions of the model then through tangential assumptions that have been conveniently added. Thus, I think a healthy skepticism on this front is warranted.

The third reason why I do not see CD theories as overarching stems from dynamic systems approaches in particular: I have some doubt as to whether the embodiment principle is always critical to cognitive development. Jordan (1972) brought up the case of normal intelligence without normal sensorimotor activity, in congenitally paralyzed individuals only able to make eye movements. Jordan described a query about this that he sent to Piaget, who forwarded the letter to H. Sinclair to make a reply. Sinclair's reply suggested that eye movements (and the mouth movements involved in eating and drinking) can be sufficient sensorimotor activity to allow intellectual development. Jordan commented on this reply as follows (p. 381):

> No matter how I tried to find this short answer satisfactory, I could not do so. I cannot reject the conclusion that this 'explanation' saves the theory nominally, at the expense of stripping it of practical significance: if the mere fact of autonomic living of the body in addition to an intact sensorium is sufficient for the constituting of sensory-motor schemata necessary for the subsequent development of adult intelligence, then all the other sensory-motor schemata constituted by the action of a normal body, upon which Piaget restricted his research and based his theory, cease to be *necessary* conditions for the development of intelligence, no matter what demonstrable effect they may have upon this development.

It would seem that a more potent factor is the ability to act upon the environment. Segalowitz (1980) commented that eye movements are not secondary circular reactions (in that they do not intrinsically change the environment) but, in conjunction with a willing caretaker who perceives the eye movements as communication, they can affect the environment. This may be what is necessary for intellectual development. Although it usually occurs through sensorimotor activity, the dynamic systems approach must separate what is usual in intellectual development from what is necessary and sufficient.

Given that I do not consider CD approaches to be 'overarching', where do they stand from my theoretical vantage point? Overall, I subscribe to the view of theories articulated by Albert Einstein (as quoted by Asimov & Shulman, 1988, p. 326):

> Creating a new theory is not like destroying an old barn and erecting a skyscraper in its place. It is rather like climbing a mountain, gaining new and wider views, discovering unexpected connections between our starting point and its rich environment. But the point from which we started out still exists and can be seen, although it appears smaller and forms a tiny part of our broad view gained by the mastery of the obstacles on our adventurous way up.

Einstein's theory did not obliterate Newton's theory; under ordinary circumstances, the latter is still a useful simplification. Similarly, in cognitive developmental theory, a symbolic level of analysis may be adequate for many uses and the application of dynamic principles may sometimes be superfluous.

What I envision, then, is a multi-level system of cognitive developmental theory that researchers can use to best advantage, with a CD level to be used as necessary, and also a level of the emergent cognitive faculties. At this second level, one might be concerned with speed and capacity parameters and how they develop (Cowan *et al.*, 2002) and with the development of the ability to use central executive processes or controlled attention to initiate and carry out strategies (Bayliss, Jarrold, Gunn & Baddeley, in press; Engle, Kane & Tuholski, 1999). In a task analysis, these parameters and abilities would be considered together and often may suffice to predict whether a child will succeed or fail at a task (Halford *et al.*, 1998). Whenever a logical processing analysis cannot explain results, the CD level of analysis would be available to offer further insights.

A key issue: the focus of attention and awareness and its development

The CD approach might naturally describe part of the human processing system and its development, but not the entire system as it is currently formulated. If the CD approach is applicable to the entire system, additional work will be needed for the portion that perhaps has been addressed least: the focus of attention and awareness and the voluntarily influenced strategic processing that seems to depend on it. Such processes have been addressed recently in the literature, in ways that may be applicable to CD approaches. Thus, attention-related phenomena might constitute a tractable future challenge for CD approaches.

Automatic and attentional components of the processing system

Cognitive psychologists distinguish between the diverse information that is processed at once in the human being and a smaller amount in the focus of an individual's awareness or attention. Similarly, Piaget (1974/1976), writing about the grasp of consciousness, opened with a demonstration that many adults incorrectly describe an act as fundamental as crawling on all four limbs, even after trying it out. Controlled processing involves more effort and attention than automatic processing, but it also is much more accessible to awareness and is much easier to stop mid-stream and to remember (Shiffrin, 1988).

The components of working memory also appear to map onto the distinction between attention-demanding and attention-free processing. Cowan (1988, 1995, 1999) conceived of working memory as including a broad field of temporarily activated memory elements that could easily be seen as a parallel network, but including also a subset of that activation, the focus of attention, that holds only a small amount of information at once and allows it to be processed further. The attentional focus is controlled partly by voluntary processes and partly by involuntary orienting to changes in the environment. Many studies demonstrated that unattended representations can be active concurrently but are lost from memory quickly; whereas Cowan's (2001) review showed that the subset of activated information that is attended appears to have a capacity limit of about four independent chunks of information in adults (and fewer in children), a parameter that surely is relevant to task performance.

Automatic and attentional processes in development

Within this view of working memory and attention, how can one conceptualize the development of automatic and attentional processes? Cowan *et al.* (2002) provided evidence that the values of several different parameters of information processing change in childhood: (1) the persistence of information that is in an activated state, (2) the capacity of the focus of attention and (3) the rate at which information can be transferred from activated memory to the focus of attention. Although it is unclear from this research if all three are completely separate as opposed to correlated changes, it is clear that they are not three different measures of the same, underlying developmental change.

Evidence that separate developmental changes occur in automatic and attentional processing comes from work in my laboratory with memory for lists of digits that are unattended (or, at best, very poorly attended) when they are presented. In the basic procedure, participants ignore most lists as they play a silent computer game involving rhymes. Occasionally (once per minute or so), the computer game is interrupted by a response grid, at which point participants must use the keyboard to recall the last list that was presented. To do this, the participant must focus attention on the sensory memory of the list before sensory memory is lost. (This method allows for the extraction of information from memory but curtails rehearsal during list presentation, so that the items stay separate in memory.) Cowan, Nugent, Elliott, Ponomarev and Saults (1999) varied the length of the list and always used a very short retention interval between the end of the last list and the response grid. The number of items that could be recalled in the correct serial positions was constant across list lengths but increased with development (from about 2.5 items in first-grade children to about 3.5 items in adults). This would appear to reflect an increasing attentional capacity, applied to the stimulus stream only after the presentation of the response cue.

Using a similar procedure but only one list length (span length), Cowan, Nugent, Elliott and Saults (2000) used retention intervals of 1, 5 or 10 s, defined as the duration of continuation of the rhyming game between the end of the target list and the onset of the response grid. Under these conditions, the rate of loss of information for the list as a whole was the same across age groups. However, for the final serial position of the list, the rate of loss was much faster for second-grade than for fifth-grade children. This result suggests that the only developmental difference in the rate of memory persistence is for uninterrupted sensory memory of the last item, not for the entire temporal stream of memory for ignored sounds.

It is my belief that such results provide important constraints that must be incorporated into any attentional model, including CD models. Are automatic and attention-demanding aspects of working memory interchangeable or can developmental changes in various tasks be modeled (and hopefully predicted) better using one type of change or another? Using attentional manipulations may be an important step in determining whether the right kind of mechanism has been incorporated into the model.

This developmental parameter-changing view raises many questions regarding possible CD models of attention. How do different aspects of CD models map onto these parameters? How can they incorporate notions of a limited-capacity working memory (presumably based on the focus of attention), as exemplified by the work on relational complexity by Halford *et al.* (1998)? CD models may be in need of further specification so that they can be definitively mapped onto these kinds of parameters.

The model of the A-not-B error described by Spencer and Schöner (this issue) appears to be a case in point. Working memory is represented as a field. It maintains knowledge of the last place that the object was hidden (on critical trials, Container B) and must compete with information on what response was made previously (removal of Container A). With development, the working memory field grows more persistent and is better able to compete; but how does that developmental change occur? It is stated that 'underlying the performance of each field is a local excitation/lateral inhibition function' and that 'a small change in parameters of the model can lead to qualitatively different behaviors over development'. In principle, though, at least three underlying changes could lead to the developmental difference, according to Cowan *et al.* (2002). Neural activation could automatically last longer in older children; the older children could use attention more efficiently to preserve the relevant information or to inhibit the activation of irrelevant information; or the older children could more rapidly and efficiently shift information from its automatically activated form into the focus of attention, before it can decay (e.g. they could have a better rehearsal routine). The Spencer and Schöner model assumes that the mechanism that operates to strengthen working memory with development early in infancy is the same one that continues to improve in childhood, but that may not be the case. For example, neural changes could lead to an automatic change in the persistence of activation early in development, followed by an improved ability to use attention later in development. A useful exercise for the CD approaches is to ask which parameters operate automatically and which operate only with the involvement of attention, and to make predictions for developmental change in the role of attention.

The use of attention in long-term learning also is in need of clarification. Munakata and McClelland (this issue) note that 'in connectionist models, learning typically occurs gradually, as small changes are made to connection weights'. This seems fine for many types of development but it is unclear whether it is fine to model the one-trial learning that often occurs. It seems apparent that one-trial learning depends on attention, but how? Is the information of sufficient importance and clarity simply stamped into the memory system, seemingly contradicting the connectionist mechanism; or is it recalled because it is rehearsed many times over, allowing time for connectionist learning to take place on the basis of covert stimulation?

Awareness, the focus of attention, and the brain

Although there are many open questions regarding how effectively CD approaches can handle the development of attention and working memory, recent approaches to the neural bases of attention do seem CD-friendly. For instance, awareness has been addressed by O'Brien and Opie (1999). From a connectionist standpoint, they advocated the viability of a 'vehicle' theory of consciousness and contrasted it with a 'process' theory that they attributed to most cognitive psychologists. The idea of a vehicle theory is that concepts that are explicitly represented in the brain, through activated neural circuits, are available to consciousness (i.e. that active thought is the vehicle for consciousness). Many other concepts are latent in the brain but are not currently active; those concepts influence thinking covertly, but are not part of consciousness. This seems in keeping with how Munakata and McClelland (this issue) also conceive of the issue, accounting for dissociations between action and awareness of action with the notion that awareness may require a stronger representation than the action itself. In contrast, a process theory stipulates that the activation of special neural circuits, not just those that represent a concept, is required for conscious awareness of the concept.

As I have argued previously (using different terminology), I believe that a process approach is correct because some areas of the brain (e.g. some parietal areas) appear to be essential for conscious thought, but not for unconscious thought (Cowan, 1995). However, I also believe that a process approach can be consistent with CD theory, notwithstanding the discussion by O'Brien and Opie (1999). Consider a specific example – object perception – for which attention is critical. Different types of features of the environment (color, shape, sound and so on) are processed by at least partly separate neural circuits. When multiple objects are present at once, there has to be a way for the brain to know how to recombine features correctly into objects. For example, given a red circle and a blue square, how does the brain figure out that the redness goes with the circle and the blueness with the square? When the combinations are arbitrary, the brain cannot solve the problem for all objects at once; in order to find a blue square in a field with other blue objects and other colored squares, the field must be inspected slowly, for an amount of time that grows in proportion to the number of objects present (Treisman & Gelade, 1980). It appears that only a subset of the field enters conscious awareness at a time under these circumstances, with the correct combination of features limited to that subset. To me, this phenomenon appears to support a process approach to consciousness.

In a neural modeling approach that seems in keeping with process theories of consciousness, but still compatible with CD theories, it has been proposed that temporal synchrony carries the relevant information for consciousness. Neurophysiological researchers have detected an approximately 40-Hz brain rhythm that seems to pulse in synchrony for feature detectors representing different features of the same object, but out of synchrony for detectors representing features of different objects (Gray, König, Engel & Singer, 1989; for a review see Cowan, 2001). Moreover, an enhanced 40-Hz rhythm for attended objects has been detected in adult humans (e.g. Tiitinen, Sinkkonen, Reinikainen, Alho, Lavikainen & Näätänen, 1993). The theory is that conscious awareness of objects depends on the synchronized pulses (necessarily including activity in some special brain areas, I would maintain; see Cowan, 1995).

If temporal synchrony plays a key role in binding object properties, CD models may not only be able to account for properties of object perception; they might also help explain the origin of capacity limits. In particular, Lisman and Idiart (1995) described a model of working memory in which different cycles of the 40-Hz rhythm carry information about different items. Putting each item on a separate cycle prevented them from being confused with one another. A capacity limit was said to emerge because the 40-Hz neural rhythm was carried within a slower frequency and all items in working memory had to be repeated in each cycle of this slower rhythm. This theory was used to justify a 7-item limit but the known parametric range just as easily could justify a 4-item limit (Cowan, 2001). Moreover, developmental change in the capacity of the focus of attention theoretically could result from neural changes that alter the relation between the 40-Hz rhythm (and its frequency) with the slower rhythm upon which it is carried; according to the approach of Lisman and Idiart (1995), neural timing differences with development would translate into attentional capacity development. (For another promising CD-type model of capacity limits, see Usher, Haarmann, Cohen & Horn, 2001.) This idea of attentional activity standing out against a busy background of neural activity seems akin to the CD approach, and may offer exciting possibilities for the future.

Concluding remarks

The CD approach seems not to be an overarching theory for everyone but is a very useful level of theoretical analysis. In my opinion, its impact will depend partly on what an 'outsider' (such as myself) actually can use. (I would not want a stark choice of either admiring the approach from the sidelines or giving up all worldly goods to follow it as a new cult!) Basic principles, such as the influence of multiple fields on performance, can be widely used. Beyond that, in the long run, it would help to construct programs that allow a relatively non-mathematical individual quickly to tamper with parameters and crank out predictions from CD models. It would be essential that these predictions be accompanied by hints as to what parameters or factors are accounting for the form of the results given particular parameter settings. When individuals who are not committed to a particular theoretical approach have the capability of extracting a priori predictions, the chances of disconfirming aspects of the theory should increase.

Attention and awareness constitute important future challenges for CD approaches. CD approaches tend to emphasize parallel processing and self-organizing, emergent systems. Thus, their limits can be tested by seeing how effectively one can account for data supporting a distinction between automatic processes, on one hand, and effortful or attention-demanding processes, on the other hand. CD approaches have a bright future, but, at this point, they are clearly not all things to all people.

Acknowledgements

I thank Jana Iverson, John Spencer and an anonymous reviewer for helpful comments. This article was completed with support from NIH Grant R01 HD-21338.

References

Asimov, I., & Shulman, J.A. (1988). *Isaac Asimov's book of science and nature quotations*. New York: Weidenfeld & Nicolson.

Bayliss, D., Jarrold, C., Gunn, D., & Baddeley, A. (in press). The complexities of complex span: explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General*.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, **104**, 163–191.

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford Psychology Series No. 26. New York: Oxford University Press.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge: Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, **24**, 87–185.

Cowan, N., Elliott, E.M., & Saults, J.S. (2002). The search for what is fundamental in the development of working memory. In R. Kail & H. Reese (Eds.), *Advances in Child Development and Behavior*, **29**, 1–49.

Cowan, N., Nugent, L.D., Elliott, E.M., Ponomarev, I., & Saults, J.S. (1999). The role of attention in the development of short-term memory: age differences in the verbal span of apprehension. *Child Development*, **70**, 1082–1097.

Cowan, N., Nugent, L.D., Elliott, E.M., & Saults, J.S. (2000). Persistence of memory for ignored lists of digits: areas of developmental constancy and change. *Journal of Experimental Child Psychology*, **76**, 151–172.

Engle, R.W., Kane, M.J., & Tuholski, S.W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge: Cambridge University Press.

Gray, C.M., König, P., Engel, A.K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, **338**, 334–337.

Halford, G.S., Wilson, W.H., & Phillips, S. (1998). Processing capacity defined by relational complexity: implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, **21**, 723–802.

Jordan, N. (1972). Is there an Achilles' heel in Piaget's theorizing? *Human Development*, **15**, 379–382.

Lisman, J.E., & Idiart, M.A.P. (1995). Storage of 7 ± 2 short-term memories in oscillatory subcycles. *Science*, **267**, 1512–1515.

Massaro, D.W. (1988). Some criticisms of connectionist models of human performance. *Journal of Memory & Language*, **27**, 213–234.

O'Brien, G., & Opie, J. (1999). A connectionist theory of phenomenal experience. *Behavioral and Brain Sciences*, **22**, 127–196.

Piaget, J. (1976). *The grasp of consciousness. Action and concept in the young child*. Translated by S. Wedgwood. Cambridge, MA: Harvard University Press. Originally published in French in 1974.

Segalowitz, S.J. (1980). Piaget's Achilles' heel: a safe soft spot? *Human Development*, **23**, 137–140.

Shiffrin, R.M. (1988). Attention. In R.C. Atkinson, R.J. Herrnstein, G. Lindzey & R.D. Luce (Eds.), *Stevens' handbook of experimental psychology* (Vol. 2) (pp. 739–811). New York: Wiley.

Tiitinen, H., Sinkkonen, J., Reinikainen, K., Alho, K., Lavikainen, J., & Näätänen, R. (1993). Selective attention enhances the auditory 40-Hz transient response in humans. *Nature*, **364**, 59–60.

Treisman, A.M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97–136.

Usher, M., Haarmann, H., Cohen, J.D., & Horn, D. (2001). Neural mechanism for the magical number 4: competitive interactions and non-linear oscillations. *Behavioral and Brain Sciences*, **24**, 151–152.