The Development of Working Memory

John P. Spencer [ID]
School of Psychology, University of East Anglia

Abstract

Working memory is a central cognitive system that plays a key role in development, with working memory capacity and speed of processing increasing as children move from infancy through adolescence. Here, I focus on two questions: What neural processes underlie working memory, and how do these processes change over development? Answers to these questions lie in computer simulations of neural-network models that shed light on how development happens. These models open up new avenues for optimizing clinical interventions aimed at boosting the working memory abilities of at-risk infants.

Working memory is a cognitive system that actively holds information in mind to facilitate cognitive operations. For instance, working memory would play a key role if you were asked to type the first, fourth, and fifth digits of a personal identification number (PIN) into a banking app. In this case, you must "load" information into an active, working memory state from long-term memory; hold that information in mind for a short duration; and then manipulate the information to achieve the goal. Note that PINs are kept to four to six digits for a reason—any longer and the number would exceed the limited capacity of working memory.

Working memory plays an important role in child development, and many theories of cognitive development take increases in working memory capacity as a starting point for cognitive change (Case, 1985). For instance, increases in working memory capacity are thought to underlie improvements in speed of processing as well as improvements in children's reasoning ability (Kail, 1991). More generally, individual differences in working memory abilities correlate with measures of children's academic performance and general intelligence (Conway, Kane, & Engle, 2003). Given these findings, it is not surprising that deficits in working memory are thought to play a central role in neurodevelopmental disorders such as attention-deficit/hyperactivity disorder (Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005).

Recent studies suggest that working memory is open to intervention (Diamond, Barnett, Thomas, & Munro, 2007). This raises exciting potential to overcome the limited working memory abilities of at-risk children (Vicari, Caravale, Carlesimo, Casadei, & Allemand, 2004). Nevertheless, questions have been raised about whether working memory training extends outside the laboratory to affect how children deploy working memory in real-world settings (Diamond & Lee, 2011). One possible way to boost the effectiveness of interventions is to intervene early in development. This would capitalize on the massive brain plasticity evident in the first few years of life.

Work in this direction has shown promise. For instance, early interventions designed to encourage caregivers to notice what their infants are attending to and "follow" their infants into those attentional episodes have had positive impacts on the cognitive outcomes of at-risk children (Landry, Smith, & Swank, 2006). Episodes of joint attention are thought to enhance working memory because they keep infants focused on the same object for longer, improving infants' representation of the object and allowing the caregiver to contingently respond to infants' bids for information. Although caregiver-based interventions are promising, their effectiveness has been limited by such factors as individual differences in caregivers' abilities, social-contextual differences, and stressors in the home.

It might be possible to optimize such interventions if we understood *how* working memory develops in

Corresponding Author:
John P. Spencer, University of East Anglia, School of Psychology
E-mail: J.Spencer@uea.ac.uk

infancy and early childhood. The present article focuses on this question and on how neural-network models have shed light on the processes that underlie changes in working memory capacity. Moreover, I highlight the advantages of having a model that researchers can use to test developmental hypotheses about how working memory changes as well as how to intervene in cases in which children are at risk.

What Neural Processes Underlie Working Memory?

Efforts to understand how neural activity gives rise to working memory have focused on sustained activation (Constantinidis & Steinmetz, 1996; Fuster & Alexander, 1971; Miller, Erickson, & Desimone, 1996). In particular, when information is held in working memory, neural populations go into a self-sustaining state in which activation persists even in the absence of an external stimulus. For instance, consider a situation in which a participant is shown an object at a spatial location for 1 s; the object is hidden; and then 5 s later, the participant is asked to point to the remembered location. When the participant sees the initial object, the object's location is encoded by the brain, creating an increase in neural activity in parts of the brain sensitive to spatial information. This activity persists during the delay as if the brain is re-presenting the stimulus to itself. After the delay, the brain then translates this remembered location into a pointing response to the location in space.

To understand how self-sustaining activation is realized in the brain, researchers can build artificial neural networks that simulate the millisecond-by-millisecond processes hypothesized to underlie sustained activity. One way to do this is to build detailed models of individual neurons and connect them into a more complex network (Deco, Rolls, & Horwitz, 2004). Such models are good at simulating how brains work in detail; however, they are not so good at explaining how people work, because they fail to explain data from behavioral experiments. For instance, sometimes people forget locations, whereas other times they make large but systematic errors (Schutte & Spencer, 2009). To understand these errors, we need neural-network models that can do what people do.

An alternative is to use models that simulate how populations or groups of neurons work without worrying about the details of each individual neuron. There is a long history of this type of approach in psychology and neuroscience. For instance, connectionist models adopt this type of simplification (e.g., Rumelhart, McClelland, & the PDP Research Group, 1986). These simpler models can be more readily matched up with behavioral data, but, importantly, they still simulate what the brain does (Bastian, Schöner, & Riehle, 2003; O'Reilly, Braver, & Cohen, 1999).

Dynamic field models are an example from this class of neural-network models (Schöner, Spencer, & the DFT Research Group, 2016; Spencer & Schöner, 2003). Dynamic field models simulate the activity of neural populations from millisecond to millisecond as the model engages in a task. For instance, a researcher might show the model a location, have it remember the location, and then "point" to the location after a 5-s delay. Such models have proven quite effective at capturing both self-sustaining brain activity as well as the interesting errors children and adults make in such simple tasks (Schutte & Spencer, 2009).

Note that dynamic field models have most often been applied to spatial and visual working memory, and the remainder of this article will focus on these examples. Readers may be familiar with proposals within this domain that working memory consists of a limited number of fixed-resolution representations in independent memory "slots" (Luck & Vogel, 1997). An alternative view holds that visual working memory is better conceived of as a shared resource that can be flexibly distributed among the visual items (Bays & Husain, 2008). The ideas presented here differ in that I take an explicitly neural approach to understanding how the brain might implement working memory. Nevertheless, there are interesting links to these other views. For instance, working memory representations in a dynamic field model have an "all-or-none" quality, similar to notions that working memory slots are either occupied or not (for a detailed discussion of these parallels, see Johnson, Simmering, & Buss, 2014).

It is also important to acknowledge that visuospatial working memory is only one corner of the working memory world. For instance, a large body of research has examined the behavioral and neural bases of phonological and verbal working memory (for a review, see, e.g., Acheson & MacDonald, 2009). Although many of the concepts used here are applicable to this work, new concepts are also needed, such as how information can be learned in sequence. Dynamic field solutions to these issues have been proposed (Sandamirskaya & Schöner, 2010), but it remains an open challenge to extend the models described herein to this research.

Self-Sustaining Activation as the Neural Basis of Working Memory

Figure 1a shows an example of a dynamic field model in a spatial working memory task in which participants are asked to remember the location of each object.[1] One second after the start of the simulation, the model

was shown three objects distributed along the feature space (*x*-axis)—a left object, a center object, and a right object. Neural sites in the model are ordered along the *x*-axis according to a functional topography, meaning neurons that "prefer" or are sensitive to leftward locations are on the left side of the axis, whereas neurons that prefer rightward locations are on the right side of the axis (for a discussion, see Schöner et al., 2016).

In response to these inputs, the model built three activation peaks—the three bumps with red on top (red indicates more intense brain activity; see *z*-axis). The three objects were visible for 2 s and then hidden for the remaining 5-s delay period (*y*-axis). As can be seen in Figure 1, the right peak continued to be active through the delay (indicated by the red arrow), whereas the left and center peaks decayed away.

Figure 1b shows the state of the dynamic field at the end of the simulation to clarify this. The dark blue line shows the activation across the field, that is, across all spatial positions from 0 to 100, and the height shows the activation intensity (*y*-axis). Note that there is only one peak on the right; activation around the left and center locations is inhibited.

Why does the peak on the right self-sustain, and why do the peaks on the left and center decay away? To answer these questions, I need to explain the rules that govern how activation evolves in the model from millisecond to millisecond.[2] The first thing that influences the activation level at each site in the field (i.e., at each spatial position) is the external input to the model. Much like in an experiment, I can turn inputs on, boosting the model's activation level at the left, center, and right locations.

What makes dynamic field models—and cortical fields in the brain—really interesting is that they can be more than just input-driven devices. When activation at, say, the center location (50) grows in response to input, something special happens as activation goes above 0—Neuron 50 starts to "talk" to its local neighbors. The rule specifying how neurons talk to each other is shown in Figure 1c. This function specifies that when Neuron 50 is active, it excites its local neighbors (e.g., neighbors ±10 locations away) and inhibits its neighbors far away (see the negative dips under the dashed blue line). The function in Figure 1c is called a *local-excitation/surround-inhibition function*, a common form of neural interaction in the brain (Fuster & Alexander, 1971).

Why is this function so important? It keeps activation going even when input is removed. In particular, Neuron 50 activates Neuron 51; reversely, Neuron 51 activates neuron 50.[3] If this reciprocal or *recurrent* activation is strong enough, the local peak maintains itself even when the input is removed. Importantly,

surround inhibition keeps the activation contained, ensuring that excitation does not spread across the entire field (causing a seizure). In summary, local excitation and surround inhibition conspire to keep the right peak in Figure 1a active during the delay.

But why do the left and center peaks die out? It turns out that there is actually nothing special about the right peak in this simulation—what is special here is that the field can maintain only one peak. If I run the simulation multiple times, sometimes the left peak will stick around, sometimes the right peak will stick around, and so on.[4] Thus, the key question is why only one peak can be maintained.

The reason is that neural interactions—specified by the interaction function in Figure 1c—are not strong enough to keep multiple peaks going. In particular, excitation is not strong enough in this case to overcome two forms of inhibition. First, neighboring peaks can kill one another because they are mutually inhibitory. As an example, look at the part of the blue line in Figure 1b inside the red oval—notice that activation is suppressed in this region. This inhibition is coming from the right peak. Given this, neurons in the region inside the red oval will be slightly less active and slightly less likely to keep local excitation going. Evidence from adults shows that such neighborhood effects exist in working memory experiments, that is, people often forget items that are close together in space (Franconeri, Jonathan, & Scimeca, 2010).

The other reason peaks die in Figure 1 is that each peak contributes a bit of global inhibition. In particular, notice that the function in Figure 1c is slightly negative to the far left and right (it is below the dashed line). This means that activated sites will excite their neighbors close by, inhibit their neighbors far away, and slightly inhibit neighbors very far away. Consequently, the peak on the right makes it a bit harder to maintain a peak anywhere else in the field. This contributes to a capacity limitation in the model—the model can maintain only a certain number of peaks (Johnson, Spencer, & Schöner, 2009).

How Does Working Memory Capacity Change Over Development?

Now that I have dissected how information is actively maintained in working memory via self-sustaining activation and the origin of capacity limits, let us turn to the really interesting question: How does working memory capacity change over development? Figures 1d and 1g show two additional simulations of the dynamic field model. Both simulations implement the same basic working memory task—present three objects and remember them for a delay—but the end result is clearly different.
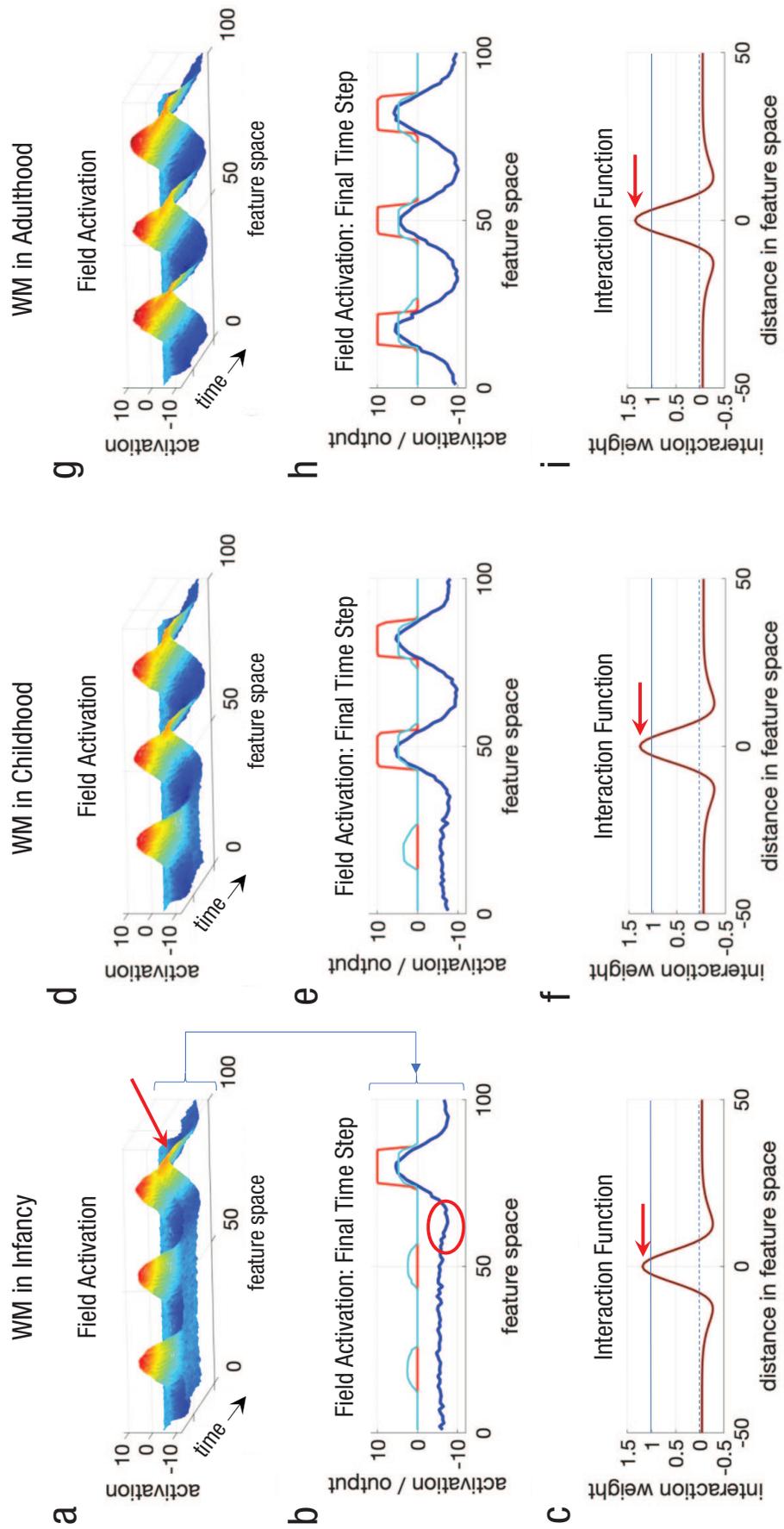
Fig. 1. Simulations of a dynamic field model showing an increase in working memory (WM) capacity over development from infancy (left column) through childhood (middle column) and into adulthood (right column) as the strength of neural interactions is increased. The graphs in the top row (a, d, g) show how activation (z-axis) evolves through time (y-axis) across the feature space of locations in the task space (x-axis). The red arrow in (a) indicates a WM peak that is maintained during the delay when no objects are visible. The panels in the middle row (b, e, h) show the state of the field activity at the last time step (indicated by brackets connecting the last time step in a to the activation profile in b). The dark blue line shows the activation level (y-axis) over feature space (x-axis), the red line shows which neurons are engaged in neural interactions (i.e., above zero activity), and the cyan line shows the strength of the long-term memory trace. The red oval in (b) indicates the region in which activation is suppressed. The graphs in the bottom row (c, f, i) show the rule governing how neurons talk to one another, with local excitation around the activated site (0) and surround inhibition to the left and right (at farther distances in feature space). The red arrows in the bottom row indicate where the strength of local excitation is increasing.

4

The model in Figure 1d remembers two items after the delay, whereas the model in Figure 1g remembers all three locations.

What did I change to create this magic? If you look closely at the panels in the bottom row—Figures 1c, 1f, and 1i—you will see that the height of the interaction function is increasing from left to right, that is, local excitation is stronger. This subtle change is all that is required to boost capacity from one item—akin to the working memory abilities of an infant (Figs. 1a–1c; see Ross-sheehy, Oakes, & Luck, 2003)—to two items—akin to the working memory abilities of a young child (Figs. 1d–1f; see Simmering, 2016)—to three items—akin to the working memory abilities of an adolescent or adult (Figs. 1g–1i; see Luck & Vogel, 1997).

This idea has been formalized in the *spatial-precision hypothesis*: Over development, children's working memory abilities improve via an increase in the strength of local excitation and surround inhibition (Schutte, Spencer, & Schöner, 2003). This simple hypothesis is actually the first formal proposal for how a neural system can increase capacity over development (Simmering, 2016). Critically, it also leads to a host of consequences. Peaks with stronger neural interactions grow faster; thus, processing speeds increase. Stronger peaks are less likely to decay; thus, forgetting decreases (Schutte & Spencer, 2002). Stronger peaks are also less likely to be pushed around; thus, memory becomes more accurate with fewer systematic biases (Schutte & Spencer, 2009). Peaks with stronger interactions are narrower; thus, discrimination improves (Simmering & Spencer, 2008). Finally, capacity increases as more peaks can be actively maintained simultaneously (Simmering, 2016). That is a lot of bang for the buck!

But how does all of this happen? Clearly, there is not a controller sitting in the brain, turning up the strength of neural interactions. One possibility is that this activity reflects structural changes in the brain such as increases in myelination (for a discussion, see Edin et al., 2009). Myelin is a fatty insulating layer that surrounds neural fibers, enhancing the efficiency of neural conduction. Because myelin makes brain signals more efficient, it could effectively increase the strength of excitation and inhibition. Although this may contribute to some of the changes evident in development, this account also raises the question of what causes changes in myelination—that is, what causes the cause?

My colleagues and I have pursued an alternative possibility (see, e.g., Perone & Spencer, 2013). To explain this, I will return to one last detail from the dynamic field model—the cyan line in Figure 1b. This line shows the long-term memory trace. Long-term memory traces in dynamic field models are built from peaks of activation: A peak's emergence slowly boosts a connection weight (the cyan line) that strengthens the memory trace locally at the peak's location. Notice that the cyan bump is strongest where each working memory peak was centered. Memory traces build slowly—it takes multiple trials to build a robust memory trace. And memory traces decay very slowly—typically, the decay time is 10 to 20 times slower than the build time. Thus, memory traces implement a form of long-term memory that builds and decays dynamically with experience,[5] a bit like a dynamic prior in Bayesian approaches in psychology.

Importantly, the memory trace feeds back into working memory, subtly boosting excitation wherever the cyan line is greater than zero. This can lead to priming effects—faster reaction times for familiar items—because peaks will build faster when input matches sites with stronger long-term memory traces (Schöner et al., 2016). Note that in the simulations shown here, I sped the memory traces up quite a bit. With slower build and decay parameters, memory traces do a good job of capturing a host of behavioral effects (Lipinski, Simmering, Johnson, & Spencer, 2010). For instance, people in spatial memory experiments learn the distribution of locations they are asked to remember, and this creates a bias in working memory toward the average remembered location (Lipinski, Spencer, & Samuelson, 2010).

Memory traces can also create changes in working memory capacity. This is shown in Figure 2. The graphs in the left column are the same as the starting points in Figure 1, but note the red arrow in Figure 2b—this highlights the low memory-trace strength at the start of the simulation. The graphs in the middle column show the model's behavior after I gave it 20 trials of experience and directed it to remember three randomly selected locations on each trial. Critically, these locations were distributed across the spatial dimension (the *x*-axis), effectively smearing the memory trace out over many different locations. As can be seen in the graphs in the middle column, when I showed it the original three objects, it actively remembered two locations instead of just one. Note that I did not change anything about the model—I simply gave it lots of distributed experience—and the model developed a capacity of two items. This occurs because the memory trace is elevated across most of the field, effectively mimicking the increase in excitation I created in Figure 1 by increasing the strength of local excitation via the interaction function. Now that increase in excitation has emerged from the model's own experience.

The graphs in the right column show what happened when I gave the model even more experience—it then had a working memory capacity of three items. The arrow in Figure 1h again highlights the memory-trace strength, showing how this boosts excitation across all spatial locations. Thus, distributed experience remembering object
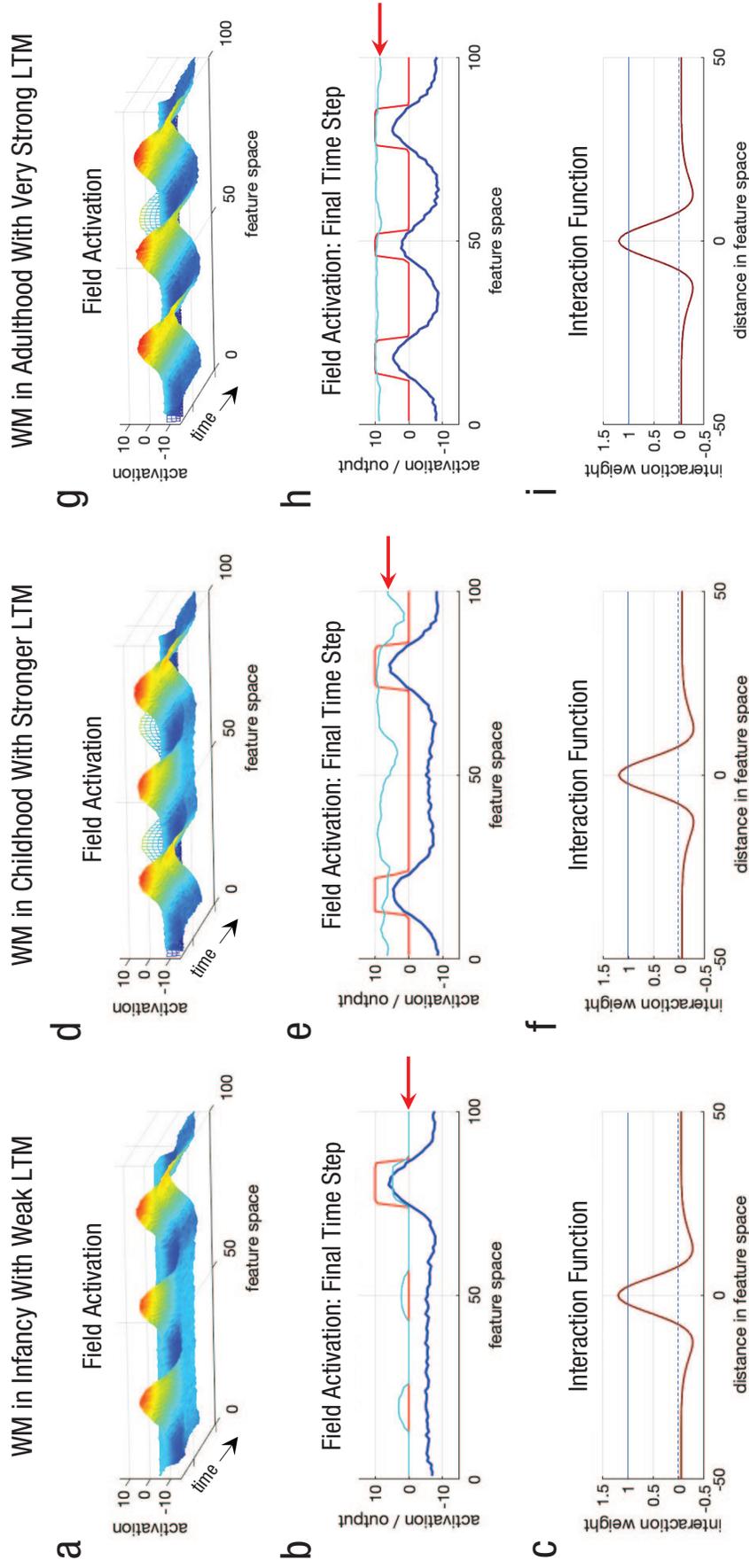
Fig. 2. Simulations of a dynamic field model showing how an increase in working memory (WM) capacity can emerge from the build-up of long-term memory (LTM) via generalized experience from infancy (left column) through childhood (middle column) and into adulthood (right column). The graphs in the top row (a, d, g) show how activation (z-axis) evolves through time (y-axis) across the feature space of locations in the task space (x-axis). The panels in the middle row (b, e, h) show the state of the field activity at the last time step. The dark blue line shows the activation level (y-axis) over feature space (x-axis), the red line shows which neurons are engaged in neural interactions (i.e., above zero activity), and the cyan line shows the strength of the LTM trace. The red arrows highlight that developmental change emerges from an increase in the strength of the memory trace across trials. The graphs in the bottom row (c, f, i) show the rule governing how neurons talk to one another, with local excitation around the activated site (0) and surround inhibition to the left and right (at farther distances in feature space). Note that the interaction functions in c, f, and i are identical.

locations in space leads to an emergent increase in spatial working memory capacity.

It is important to highlight that *distributed* experience is critical. If I simply showed the model the same three locations over and over, it would develop a great working memory for those three locations, but peaks associated with other locations would collapse because there would be only a weak excitatory long-term memory boost at those sites. Along continuous dimensions such as direction, distance, color, and orientation, distributed experience is likely the norm. Thus, these simulations point toward a general mechanism for how the accumulation of experience can drive developmental changes in working memory capacity.

Let me also emphasize that this is just a toy example given that I have made the memory traces unrealistically fast. My colleague and I have shown, however, that this basic explanation does, in fact, quantitatively model longitudinal developmental data in infancy (Perone & Spencer, 2013). In that simulation experiment, we gave our model 300,000 time steps of visual exploratory experience to simulate a few months of developmental time. This was sufficient to simulate developmental changes in visual processing speed. Decay was present in this model, but it was very slow. This enabled the long-term memory traces to gradually build over several months of experience.

Notably, we also showed in this simulation experiment that we could simulate individual differences in development. In particular, we initialized our model with very weak neural interactions, which did a good job of capturing the behaviors of preterm infants. Importantly, the preterm model developed more slowly even when given the same chances to explore the visual world. This reproduced longitudinal data from preterm infants.

Finally, we implemented an intervention in the model: After 150,000 time steps, we helped the preterm model by occasionally giving it a bit more input at whatever object location it happened to be looking (Perone & Spencer, 2013). Conceptually, this is similar to approaches used in intervention studies that have trained caregivers to follow what their infants are looking at and to help them sustain looking to these objects by holding the object, talking about the object, and so on. After several months of the intervention, the dynamic field model showed an advancement in speed of processing that was closer to the behavior of our term model—the intervention worked.

Conclusions

Working memory is a central cognitive system that changes dramatically over development with far-reaching consequences. In this article, I have highlighted work that integrates neural and behavioral findings using a type of neural-network model. This modeling work explains how the brain implements sustained activation, how these mechanisms give rise to capacity limits, and how capacity limits change over development as generalized experience accumulates. Critically, this account has explained a host of behavioral and neural findings, including several novel predictions (Johnson, Spencer, Luck, & Schöner, 2009).

What does the future hold? At present, my colleagues and I are pursuing two key issues. The first is to better understand the neural changes that underlie the early development of working memory. On this front, we have recently proposed ways to simulate brain data directly from dynamic field models, opening up tests of this modeling framework using functional MRI (Wijeakumar, Ambrose, Spencer, & Curtu, 2017) and, early in development, functional near-infrared spectroscopy (Wijeakumar, Kumar, Reyes, Tiwari, & Spencer, 2019).

The second direction is to capitalize on the promise of using neurocomputational models as intervention tools. Here, a first step is to understand how working memory measured in the laboratory relates to how working memory is deployed in the real world in, for instance, dyadic exchanges between caregiver and infant. Although this work is still in its early stages, we hope to use such information to guide and optimize dyadic clinical interventions.

For instance, in newer work, my colleagues and I developed a dyadic dynamic field model in which we can simulate both the caregiver and the infant because the models share the same virtual world (Perone, Aneja, & Spencer, 2020). This has allowed us to tune the caregiver model to match the working memory abilities of the caregiver and tune the infant model to match the working memory abilities of the infant. We can then simulate candidate interventions. What if this caregiver helped the infant dwell on objects by picking them up routinely and shaking them? Alternatively, perhaps the infant needs help releasing fixation after a minute or two of sustained attention to an object. And might the caregiver benefit from a visual or auditory reminder to help support his or her working memory abilities? An advantage of having a model is that we can implement all of these candidate interventions to see which ones lead to the best outcomes, that is, the most robust working memory development for the child.

The best candidate intervention could then be implemented, with the model making step-by-step predictions about how the intervention should change working memory over time. These constitute predictions that can be tested as the intervention proceeds. In this way, one can intervene dynamically rather than implementing a

one-size-fits-all intervention that can be evaluated only after the intervention is complete. We think this could revolutionize clinical interventions with young children.

Recommended Reading

Cowan, N. (2016). Working memory maturation: Can we get at the essence of cognitive growth? *Perspectives on Psychological Science*, *11*, 239–264. A summary of current issues in the study of working memory development.

Perone, S., & Spencer, J. P. (2013). (See References). An article in which many of the key ideas discussed in the current article were established.

Simmering, V. R. (2016). (See References). An excellent review of the development of working memory followed by novel empirical studies of development and modeling using dynamic field theory.

Spencer, J. P., Austin, A., & Schutte, A. R. (2012). Contributions of dynamic systems theory to cognitive development. *Cognitive Development*, *27*, 401–418. An overview of dynamic field theory and applications to cognitive development.

Spencer, J. P., & Schöner, G. (2003). (See References). A summary of key justifications for a dynamical systems approach to cognition using dynamic field theory.

Transparency

Action Editor: Robert L. Goldstone
Editor: Robert L. Goldstone
Declaration of Conflicting Interests
   The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

ORCID iD

John P. Spencer 🄳 https://orcid.org/0000-0002-7320-144X

Notes

1. Simulation code and instructional videos are available at www.dynamicfieldtheory.org.
2. By "rules," I mean the mathematical equation that specifies how the model evolves through time.
3. Note that the red line in Figure 1b indicates which neurons are above threshold (above zero) and are, therefore, sharing local excitation.
4. You can try this online at www.dynamicfieldtheory.org.
5. This implementation of long-term memory connects our theory (see, e.g., Perone & Spencer, 2013) with recent ideas regarding "active" long-term memory (e.g., Oberauer & Lin, 2017).

References

Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, *135*, 50–68.

Bastian, A., Schöner, G., & Riehle, A. (2003). Preshaping and continuous evolution of motor cortical representations during movement preparation. *European Journal of Neuroscience*, *18*, 2047–2058.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*, 851–854.

Case, R. (1985). *Intellectual development. Birth to adulthood.* New York, NY: Academic Press.

Constantinidis, C., & Steinmetz, M. A. (1996). Neuronal activity in posterior parietal area 7a during the delay periods of a spatial memory task. *Journal of Neurophysiology*, *76*, 1352–1355.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*, 547–552.

Deco, G., Rolls, E. T., & Horwitz, B. (2004). "What" and "where" in visual working memory: A computational neurodynamical perspective for integrating fMRI and single-neuron data. *Journal of Cognitive Neuroscience*, *16*, 683–701.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, *318*, 1387–1388.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, *333*, 959–964.

Edin, F., Klingberg, T., Johansson, P., McNab, F., Tegnér, J., & Compte, A. (2009). Mechanism for top-down control of working memory capacity. *Proceedings of the National Academy of Sciences, USA*, *106*, 6802–6807.

Franconeri, S. L. L., Jonathan, S. V. V., & Scimeca, J. M. M. (2010). Tracking multiple objects is limited only by object spacing, not by speed, time, or capacity. *Psychological Science*, *21*, 920–925.

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, *173*, 652–654.

Johnson, J. S., Simmering, V. R., & Buss, A. T. (2014). Beyond slots and resources: Grounding cognitive concepts in neural dynamics. *Attention, Perception, and Psychophysics*, *76*, 1630–1654.

Johnson, J. S., Spencer, J. P., Luck, S. J., & Schöner, G. (2009). A dynamic neural field model of visual working memory and change detection. *Psychological Science*, *20*, 568–577.

Johnson, J. S., Spencer, J. P., & Schöner, G. (2009). A layered neural architecture for the consolidation, maintenance, and updating of representations in visual working memory. *Brain Research*, *1299*, 17–32.

Kail, R. (1991). Developmental changes in speed of processing during childhood and adolescence. *Psychological Bulletin*, *109*, 490–501.

Landry, S. H., Smith, K. E., & Swank, P. R. (2006). Responsive parenting: Establishing early foundations for social, communication, and independent problem-solving skills. *Developmental Psychology*, *42*, 627–642.

Lipinski, J., Simmering, V. R., Johnson, J. S., & Spencer, J. P. (2010). The role of experience in location estimation: Target

distributions shift location memory biases. *Cognition*, *115*, 147–153.

Lipinski, J., Spencer, J. P., & Samuelson, L. K. (2010). Biased feedback in spatial recall yields a violation of delta rule learning. *Psychonomic Bulletin and Review*, *17*, 581–588.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *The Journal of Neuroscience*, *16*, 5154–5167.

Oberauer, K., & Lin, H.-Y. (2017). An interference model of visual working memory. *Psychological Review*, *124*, 21–59.

O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1999). A biologically based computational model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 375–411). Cambridge, England: Cambridge University Press.

Perone, S., Aneja, P. & Spencer, J. P. (2020). *The processes that underlie synchronous caregiver-infant interactions: Insights from a dynamic field model of dyadic visual cognition*. Manuscript in preparation.

Perone, S., & Spencer, J. P. (2013). Autonomous visual exploration creates developmental change in familiarity and novelty seeking behaviors. *Frontiers in Psychology*, *4*, Article 648. doi:10.3389/fpsyg.2013.00648

Ross-sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development*, *74*, 1807–1822.

Rumelhart, D. E., & McClelland, J. L., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations*. Cambridge, MA: MIT Press.

Sandamirskaya, Y., & Schöner, G. (2010). An embodied account of serial order: How instabilities drive sequence generation. *Neural Networks*, *23*, 1164–1179.

Schöner, G., & Spencer, J. P., & the DFT Research Group. (2016). *Dynamic thinking: A primer on dynamic field theory*. New York, NY: Oxford University Press.

Schutte, A. R., & Spencer, J. P. (2002). Generalizing the dynamic field theory of the A-not-B error beyond infancy: Three-year-olds' delay- and experience-dependent location memory biases. *Child Development*, *73*, 377–404.

Schutte, A. R., & Spencer, J. P. (2009). Tests of the dynamic field theory and the spatial precision hypothesis: Capturing a qualitative developmental transition in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1698–1725.

Schutte, A. R., Spencer, J. P., & Schöner, G. (2003). Testing the dynamic field theory: Working memory for locations becomes more spatially precise over development. *Child Development*, *74*, 1393–1417.

Simmering, V. R. (2016). Working memory capacity in context: Modeling dynamic processes of behavior, memory, and development. *Monographs of the Society for Research in Child Development*, *81*(3), 7–24.

Simmering, V. R., & Spencer, J. P. (2008). Generality with specificity: The dynamic field theory generalizes across tasks and time scales. *Developmental Science*, *11*, 541–555.

Spencer, J. P., & Schöner, G. (2003). Bridging the representational gap in the dynamical systems approach to development. *Developmental Science*, *6*, 392–412.

Vicari, S., Caravale, B., Carlesimo, G. A., Casadei, A. M., & Allemand, F. (2004). Spatial working memory deficits in children at ages 3-4 who were low birth weight, preterm infants. *Neuropsychology*, *18*, 673–678.

Wijeakumar, S., Ambrose, J. P., Spencer, J. P., & Curtu, R. (2017). Model-based functional neuroimaging using dynamic neural fields: An integrative cognitive neuroscience approach. *Journal of Mathematical Psychology*, *76*, 212–235.

Wijeakumar, S., Kumar, A., Reyes, L. M. D., Tiwari, M., & Spencer, J. P. (2019). Early adversity in rural India impacts the brain networks underlying visual working memory. *Developmental Science*, *22*(5), Article e12822. doi:10.1111/desc.12822

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, *57*, 1336–1346.